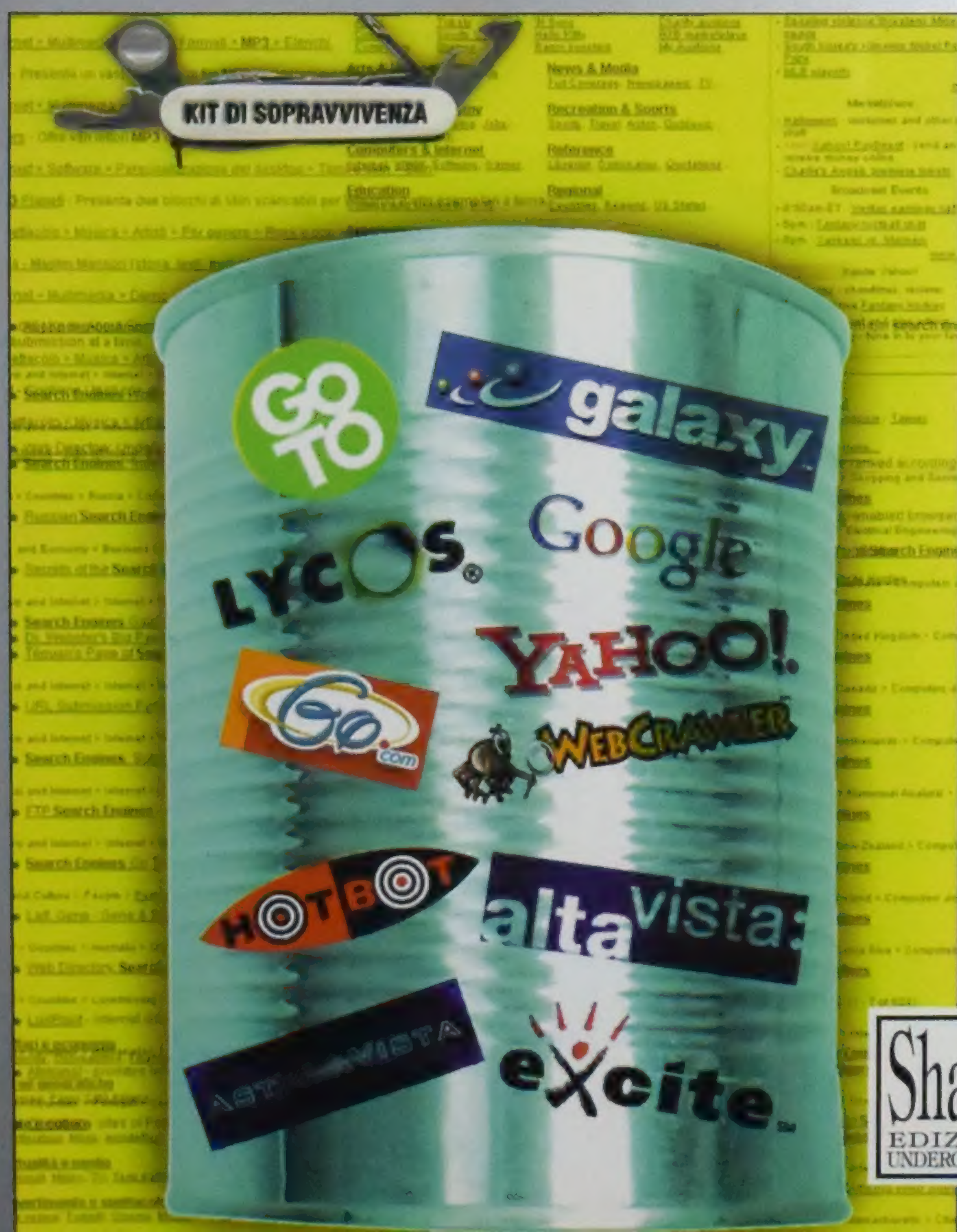
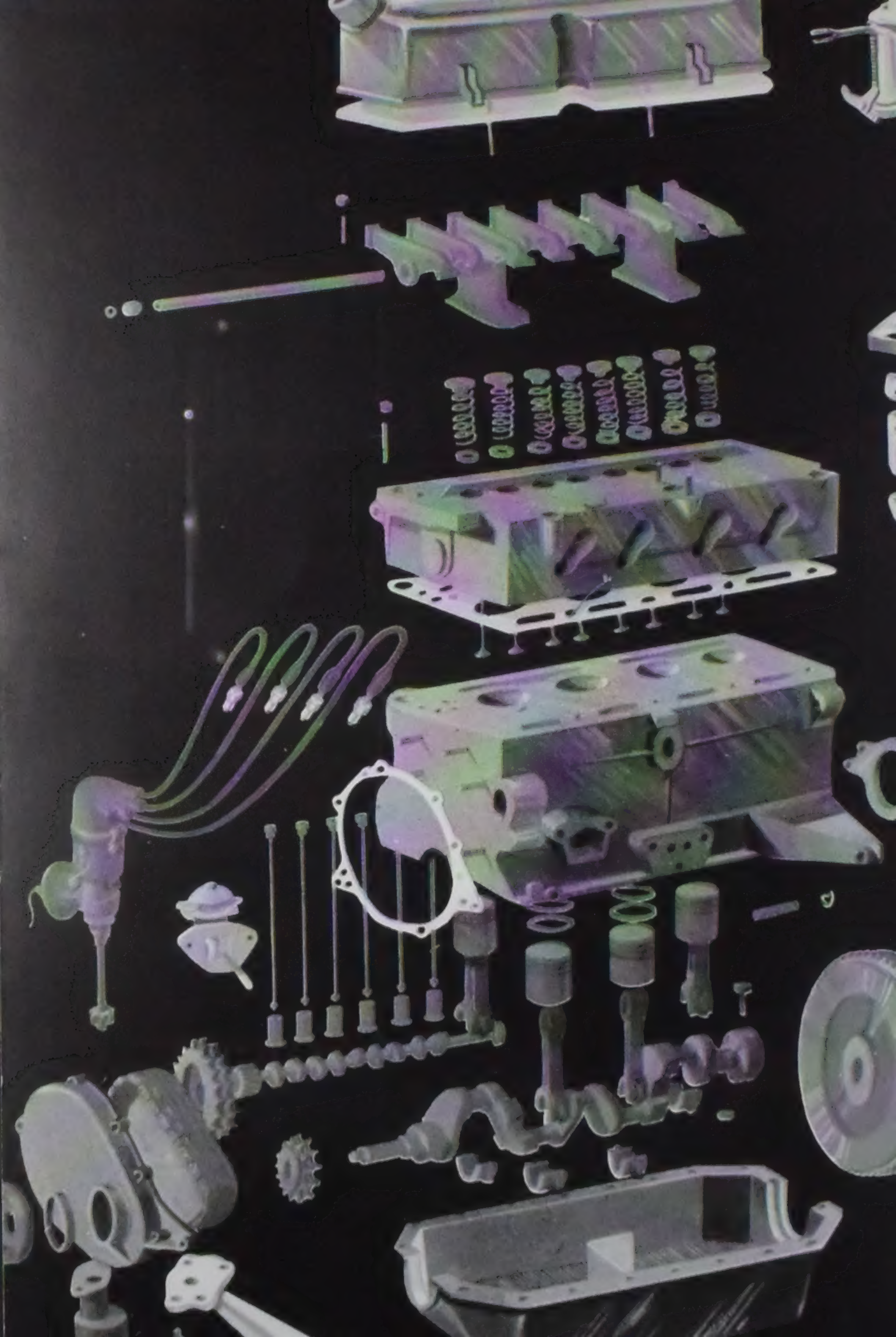
Claudio Parrini - Ferry Byte

I MOTORI DI RICERCA NEL CAOS DELLA RETE

KIT DI SOPRAVVIVENZA



ShaKe
EDIZIONI UNDERGROUND



CYBERPUNKLINE

Complicità e revisione letteraria: si ringrazia in particolare la collaborazione di Mirella Castigli per il contributo in termini di idee nella parte relativa alla "Linguistica" e nel capitolo "Caos vs. Business", e per l'indispensabile sua opera di complicità.

Grazie a: Raf Valvola, la mailing list Cyber-Rights, l'hacklab di Firenze, Shine, Francesco Galluzzi, Ubaldo Fadini, Arclele, Tommaso Tozzi, Emanuele Vecchia, Carla Maltinti, TBazz e la nostra rete di affetti tutta, per averci aiutato nella realizzazione del libro.

Editing: Sergio Cicconi

Progetto grafico: Rosie Pianeta, Paoletta Nevrosi

Immagine di copertina: Federico Bucalossi

Fotocomposizione: ShaKe Edizioni

Contatti postali: ShaKe, viale Bligny, 42 - 20136 Milano

Contatti telefonici: 02/58317306

Contatti elettronici: http://www.decoder.it
http://www.shake.it
e-mail: info@shake.it
ufficio stampa: press@shake.it

Prima edizione: 2001
Seconda edizione: marzo 2002

Stampa: Legoprint, Lavis (TN)

ISBN 88-86926-75-8

Claudio Parrini - Ferry Byte

I MOTORI DI RICERCA NEL CAOS DELLA RETE

KIT DI SOPRAVVIVENZA



SHAKE EDIZIONI UNDERGROUND

INDICE

Prefazione 7

PARTE PRIMA
L'INFORMAZIONE IN RETE

1 Il diluvio informativo 13
1.1 Informazione perenne 13
1.2 Memoria e conoscenza 16
1.3 Eccessi e censure 18
1.4 Tecniche, sapere, individuo, libertà 20

2 Organizzazione rizomatica dell'informazione in rete 25
2.1 Struttura della rete 25
2.2 Comunicazione, ovvero come mettere in comune l'informazione 26
2.3 I motori di ricerca e la rete 28
2.4 Si fa presto a dire informazione... 30

3 Centralità inedita del sapere 35
3.1 Oltre il sapere come contenuto 35
3.2 Un sapere centrale e mobile 36
3.3 Il sapere come collante sociale 37
3.4 Chi non partecipa non sa, chi non sa non può partecipare 38

PARTE SECONDA
RICERCA DELL'INFORMAZIONE IN RETE

4 Il motore di ricerca 43
4.1 Ricerca dell'informazione in rete. Comunità, agenti intelligenti, bookmark e motori 43

4.2 Funzionamento del motore di ricerca 45
4.3 Panoramica dei motori esistenti 46
4.4 Come si ricerca l'informazione in rete 48

5 Il motore di ricerca e l'informazione 53
5.1 Necessità di indicizzazione 53
5.2 Forme accessibili di informazione 55

6 Linguistica 67
6.1 Orizzonti linguistici 67
6.2 Il "significato" nei motori di ricerca 68
6.3 Nel cuore dell'indicizzazione 70
6.4 Il sapere come gioco linguistico 71
6.5 La pragmatica riformista. L'esempio del top ranking 72
6.6 La via rivoluzionaria o pratica creativa 73

PARTE TERZA
TECNICHE DI SOPRAVVIVENZA

7 Agenti software o agenti umani? 81
7.1 Limiti e prospettive degli agenti intelligenti 81
7.2 Cacciatori di informazione 84
7.3 Potenzialità e sviluppi degli strumenti collettivi di condivisione dei saperi 86

8 Filosofia dei portali 95
8.1 Tutto in una parola 95
8.2 L'importante è fidelizzare 96
8.3 Il sito grande mangia quello piccolo 97

9 Caos vs. Business? 101
9.1 Fette sempre più piccole di una torta sempre più grande 101
9.2 Ranking: ovvero come gli ultimi non saranno mai i primi 102
9.3 Strumenti di ricerca caotici come antidoti all'infinito e ai meccanismi di censura 105
9.4 Le ragioni del caos contro quelle del business nella definizione delle classifiche dei risultati di ricerca 106

Appendice manualistica 109

Bibliografia 116
Sitografia 118

PREFAZIONE

Questo libro nasce dalle discussioni, ricerche, letture ecc. che hanno attratto la nostra attenzione su un argomento così complesso e intrigante quale è quello dell'ab/uso dei motori di ricerca.

Non è retorica ma semplice verità affermare che il nostro ruolo – al di là di qualche stravagante enunciazione in merito – è stato quello di registrare sensazioni, informazioni, vibrazioni in rete (e fuori di essa) per poi (de)formarle nella scrittura di questo libro. Rendiamo quindi all'entropia di quest'universo di idee, attraverso la forma del no-copyright e destinando una versione elettronica del presente libro ad alcune directory no-profit della rete, tutti gli input di ipotesi e proposte recepite in dieci anni di attività artistica e di hacktivismo in rete.

Sì, perché entrambi – uno artista (Claudio Parrini) e l'altro hacktivista (fERRYbYTE) – appassionati navigatori della rete ben prima che i monitor diventassero più sottili dei loro computer, con una modesta preparazione tecnica ma con un bagaglio di curiosità immenso, siamo da anni impegnati a riflettere sull'effettiva incidenza degli strumenti virtuali nel reale.

Siamo venditori di pentole o agitatori di popolo?

Stiamo forse tirando il carro a chi, dalla rete, sta cercando di ricavarne profitti e quotazioni in Borsa oppure, inventandoci pratiche inedite di manifestazione come il netstrike, siamo "geni incompresi" di un'improbabile arte dell'agit-prop digitale?

Con una dose di complicità così alta, che è riscontrabile solo fra due ladri di professione costretti spesso a improvvisare per risolvere situazioni impreviste nel loro comune lavoro, e con il piacere – oppure l'illusione – di sperimentare nuove riflessioni nel magma della comunicazione digitale, ci siamo confrontati insieme per anni sul ruolo della ricerca dell'informazione in rete e sulle sue implicazioni sociologiche.

La ricerca dell'informazione in rete, e l'altra faccia della medaglia – le tecniche per rendere l'informazione più visibile nel mare magnum del Web –, sono argomenti sempre più importanti nell'attuale geografia di Internet.

L'evoluzione del Web – da frontiera per "fanatici del modem" e cowboy della consolle alla cybernavigazione di massa – è ormai sotto gli occhi di tutti. Più oscuro rimane invece il tema della gestione dell'informazione e delle risorse dell'informazione in rete: un tema che, nell'evoluzione commerciale di Internet, si può tradurre nel famoso detto baconiano "sapere è potere". E poiché i motori di ricerca sono indubbiamente una delle principali porte di accesso ai saperi disseminati in rete, è di stretta attualità analizzarne funzionamento e limiti.

I motori di ricerca, strumenti ai quali noi tutti ricorriamo a volte di fretta, acriticamente, altre volte solo con speranza o curiosità, per quanto studiati sotto il profilo tecnico da innumerevoli manuali, sono tuttora descritti come una panacea al diluvio informativo. Raramente però sono messi in discussione, o anche soltanto guardati al microscopio con occhi disincantati.

Cosa sono gli spider? Cosa vedono realmente? Cosa cercano i motori di ricerca? È possibile definire un kit di sopravvivenza per profani ma anche navigatori esperti che si affacciano su questo magma fluttuante di informazioni? Esiste un reale pericolo di "censura" dell'informazione, o c'è parità di visibilità per tutti su Internet? L'informazione è in effetti "universalmente accessibile"?

Queste sono solo alcune delle domande che ci poniamo ormai da anni nel nostro approccio sperimentale alla rete, e che ci permettono di formulare un quadro di ipotesi e suggerimenti su cui tessere la ragnatela delle nostre speculazioni.

A partire dall'analisi dell'assetto biologico della rete, abbiamo delineato un percorso originale e allo stesso tempo operativo per rispondere, da un punto di vista hacker (sperimentale e smanettone), a questi interrogativi e riuscire a districarsi nel modo più critico ed efficace possibile tra le pratiche di reperimento-visibilità-accessibilità dell'informazione in Internet.

Questo libro è un'indagine sui motori di ricerca, sulle loro caratteristiche, limiti ed evoluzioni. Un'indagine che sviscera, nella prima parte più teorica, questioni scottanti come i dibattiti sull'iperinvasione

dei dati in rete, e sulle conseguenti operazioni di filtraggio dell'informazione; ma anche un'analisi che delimita il paradigma rizomatico che caratterizza l'organizzazione dell'informazione e la messa a fuoco dell'*information retrival* nell'ambito dell'inedita centralità del sapere/know-how.

La seconda parte, più propriamente tecnica, affronta analiticamente il campo delle risorse di ricerca, spiegando in modo dettagliato il funzionamento dei motori di ricerca e delle web directory, le modalità di ricerca e tecniche del top ranking, per apparire fra le prime posizioni nei risultati di ricerca. Un capitolo è interamente dedicato alla nuova frontiera dell'accessibilità, un ruolo nodale nella costruzione del World Wide Web. Segue una trattazione di tipo linguistico-sintattico, che esamina le dinamiche di indicizzazione dei documenti sotto il profilo linguistico, a partire da concetti quali la polisemia e la pertinenza, fino al ruolo della creatività e del "gioco linguistico" applicati alla ricerca telematica.

La terza e ultima parte prende avvio dal vivace dibattito sull'impiego di agenti software (intelligenza artificiale) e personale umano nella ricerca dell'informazione, soffermandosi quindi sul fenomeno dell'anno duemila: il dilagare di portal e vortal, che spalancano un nuovo fronte teorico-critico nella mappatura dei sistemi di ricerca.

Chiude infine il libro una sezione "sperimentale", intitolata "Caos vs. Business", che è frutto delle nostre elaborazioni teoriche sull'ipotesi di motori di ricerca caotici: una freccia avvelenata di due cyber-hacktivist scagliata verso coloro che, inseguendo avidamente e senza scrupoli le rotte dell'Eldorado digitale, calpestano quei germogli che in futuro potrebbero dare origine a una vita elettronica più democratica. Ispirato alle idee più fertili tracciate dai teorici del caos nel ventennio 1960-1980, il libro propone un'intuizione artistica – di ricerca casuale – per aprire la via a una ricerca dell'informazione in rete che possa affrancarsi dai soffocanti lacci del business per recuperare, almeno in parte, le caratteristiche di una ricerca ipertestuale critica e libertaria, intimamente legata alla natura rizomatica e decentralizzata di Internet nell'epoca di Freenet e delle sottoreti autonome. Affinché non avvenga in rete quella catastrofe epistemologica e culturale che determinò la fine dell'immaginario celtico tramandato dalle antiche leggende irlandesi: quando le divinità celtiche abbandonarono all'oblio i loro popoli non appena i romani iniziarono a costruire le prime strade, seguendo i principi di Ordine e Ratio, e sostituendo con la geometrica regolarità del *castrum* la casualità dei sentieri tradizionali. Fu allora che l'immaginifico universo popolato di gnomi, folletti e magiche pozioni, di cui rimangono flebili tracce nella produzione shakespeariana e nella saga di Asterix, decise di sparire per sempre, condannando le popolazioni celtiche alla sconfitta contro l'esercito romano: il caotico mondo dei celti,

con i loro misteriosi sentieri fra boschi lunari, non poteva subire l'affronto della razionalità latina che imponeva, diritto e potenza, la costruzione di uno squadrato stradario ad angoli retti. Caos e magia cedettero così il passo a geometria e rigore...

La grande capacità di comunicazione orizzontale della rete, l'innata esigenza a fare comunità, il naturale spirito tecno-pagano di ludiche strategie antagoniste contro ogni controllo dell'informazione, riusciranno a resistere e opporsi all'assalto delle nuove potenze straniere? Piccoli gnomi della rete provano ogni giorno a respingere i tentativi di imbrigliare e monetizzare la creatività, realizzando toolkit di sopravvivenza del "Caos vs. Business". Autostrade a tripla corsia costruite sulle tracce di artigianali e storici sentieri non decreteranno la fine di paradigmi alternativi al pensiero dominante. La rete continuerà a proliferare nel suo snodarsi imprevedibile.

Questo libro, scaturito da un'esigenza intellettuale di recupero dello sperimentalismo artistico-epistemologico-linguistico nel gioco della rete, è dunque un'indagine sui motori di ricerca e sul labirinto dell'informazione, calata nel più vasto contesto delle pratiche di socializzazione delle tecnologie e condivisione dei saperi.

È un *work in progress* che, attraverso un'appendice, una selezione di siti e una bibliografia ragionata, si presenta come uno strumento dinamico in continua costruzione, sviluppandosi anche al di fuori di questo luogo cartaceo, nella deriva telematica di un sito web (http://www.strano.net/chaos) in costante aggiornamento e revisione, e soprattutto aperto ad apporti orizzontali con le mailing-list di discussione (come la storica http://www.ecn.org/lists/cyber-rights).

PARTE PRIMA

L'INFORMAZIONE IN RETE

CYBER RIGHTS

La lotta partigiana in Italia 1943-45





Il progetto di "Isole nella Rete" si propone di aprire uno spazio di visibilita' su Internet che metta in relazione il "popolo della rete"...

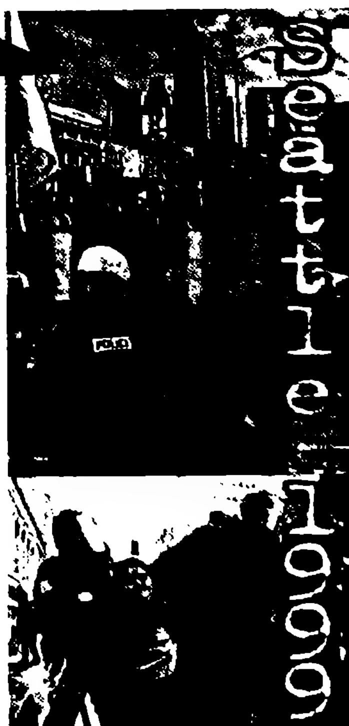






Dalla Guerra in Jug[illegible]via
ci giungevano
le immagini delle visio[illegible]
ipnotiche di Zograf,
autore serbo [illegible]rl
una corrispon[illegible]
l'Occidente [illegible]
puntuali e-mail [illegible]tti
che accadevano [illegible] lu[illegible]
"Ho provato [illegible]
una e-mail un[illegible]
ma il collegamento si
è interrotto per
un'esplosione".

CYBER RIGHTS NOW!



ezln ezln ezln ezln ezln ezln ezln ezln ezln



1

IL DILUVIO INFORMATIVO

1.1 INFORMAZIONE PERENNE

Gli sviluppi degli ultimi dieci anni di Internet, che coincidono in particolare con la storia del Word Wide Web,[1] hanno prodotto uno scenario telematico composto da un'enorme, pressoché illimitata, massa di dati informativi. Filosofi, sociologi e studiosi della rete definiscono il fenomeno come irrompente e difficilmente gestibile, nei termini di "diluvio informativo" o "diluvio informazionale".[2] La metafora adottata, che ci riporta istantaneamente alla catastrofe immane del diluvio universale, è molto efficace e puntuale. L'imponente e continuo flusso di informazioni è infatti paragonabile a una violenta pioggia ininterrotta che invece di diminuire tende sempre più ad aumentare: dati nuovi, oppure dati vecchi che giacciono da tempo in siti divenuti putridi, dati che s'intrecciano, si accavallano, si amplificano, si raddoppiano, talvolta provocando delle vere e proprie metastasi informative.

La crescita di Internet è veramente impressionante: le persone collegate aumentano sempre più in funzione di alcuni fattori nodali quali l'accesso "gratuito" alla rete, i massicci investimenti (hardware e software) da parte delle multinazionali della comunicazione e l'avvio ufficiale di sempre più estese forme di commercio online. Tutto questo non dovrebbe essere visto in termini apocalittici se gli apparati pubbli-

ci e la politica riuscissero a garantire una democrazia telematica in grado di salvaguardare i mondi individuali e le autonomie di fruizione ed elaborazione di pensiero. Anzi l'evoluzione di Internet, visto in termini di "entropia positiva",[3] può intendersi a maggior ragione come un passo accrescitivo verso la realizzazione di un intelletto collettivo umano. La libertà che Internet offre di potersi immergere nel flusso informativo, di essere in grado di pubblicare, di giudicare, di criticare non deve certo venir meno; l'importante è però non offrire sola informazione, ma far sì che questa venga compresa. È proprio quest'impegnativo compito a necessitare forme di organizzazione, gestione e diffusione dell'intero materiale che si veicola su Internet. E diventa primario anche il ruolo delle principali risorse di ricerca: i motori di ricerca, le web directory, i portali e vortali ecc.

La rete quindi può essere riassunta come un insieme di conoscenze e saperi, punti di vista e modi di pensare secondo una logica comunicativa orizzontale e provvisoria che rifiuta la gestione e diffusione gerarchica di tipo *broadcasting*, propria della logica televisiva. Data la specifica natura di Internet è necessario doverne analizzare sistematicamente le parti costitutive, per capirne e interpretarne lo spirito, l'atteggiamento e lo stile di vita di coloro che praticano il territorio telematico. Capire quindi il mezzo enucleando quegli aspetti che sono fuorvianti e deleteri in altri sistemi comunicativi: acquisendo così specifiche capacità tecniche, definendo regole e metodi di navigazione e autodifesa digitale, implementando un approccio mentale più globale e allo stesso tempo in grado di discernere dati informativi con decisioni continue e gran volontà di apprendimento.

Il caos in rete sembra tale ma non lo è: il disorientamento, l'incertezza e la probabilità sembrano piuttosto esserne le componenti necessarie al mantenimento di uno stato "fisico" ottimale. C'è un velo di apparenza nella paventata confusione telematica. La comprensibile ansia tecnologica che Internet ci scatena potrebbe portarci pertanto a due forme estreme di comportamento: da una parte un'accettazione irrazionale e indiscriminata di quanto la tecnica ci propone, dall'altra il rifiuto incondizionato, frutto di un arido "nichilismo". Sulla "questione della tecnica", abbiamo così da una parte il tecno-ottimismo utopistico alla Pierre Lévy, fondato su una teoria antropologica che pone il sapere come protagonista dello sviluppo umano; dall'altra il filone pessimistico alla Virilio e Baudrillard, che vede nelle innovazioni scientifiche una serie infinita di rischi e pericoli.

La nostra pratica quotidiana della rete intende superare queste categorie: nell'uso quotidiano della tecnologia c'è la nostra risposta. La tecnologia è un prodotto dell'uomo e come tale va compreso e utilizzato. I motori di ricerca non sono altro che delle ennesime "scatole nere" da capire; il passo successivo sta nel comunicare a tutti "il truc-

co".[4] Il percorso da compiere non deve comportare perciò alcuna riflessione inutile sulla presunta malvagità o meno di determinate tecniche; piuttosto deve mettere in moto progettualità comunitarie finalizzate al risolvimento dei problemi pratici del nostro domani. Un domani non troppo lontano, viste la velocità e la "obsolescenza" delle scoperte digitali; spesso ciò che oggi riteniamo congeniale ai nostri bisogni, in un tempo brevissimo diventa vecchio e inusabile. Le conoscenze tecniche acquisite generano un sapere che deve essere reso comunitario, con spirito mutuale e ramificato. Le mutazioni della tecnica non saranno più vissute come minacce se il sapere tecnico viene diffuso e distribuito con l'apporto di strutture statali: comunità, istituzioni, ma soprattutto la rete.

Correre ai ripari rispetto a questa perenne ipertrofia informativa usando delle difese sbagliate può essere controproducente. Per rimanere nell'ambito dei motori di ricerca, una delle difese utili ad affrontare il diluvio informativo sta nel "calibrare" e "posizionare" bene questi "congegni", in altri termini, rispondendo alla rete con la stessa moneta offerta dalla rete: l'orizzontalità. Come in tanti altri casi nella storia della cultura cyber, molti sperimentatori stravaganti e smanettoni a tutto tondo hanno dato vita a realtà poi incorporate a pieno titolo nel mondo attuale: il circuito Fidonet, l'ipertesto, Linux, i formati di compressione tipo MP3, esperienze come Gnutella e Freenet, le ricerche sulle interfacce fluttuanti. Anche noi, nel nostro piccolo ambito di artisti e ricercatori cyberattivisti, con questo lavoro, vogliamo interrogare le intersezioni possibili tra le risorse di ricerca, il continuo accumulo di informazione che Internet produce e le metodologie di ricerca: nuovi *knowbots* progettati sulle basi dell'intelligenza artificiale, e le futuribili ipotesi tecnologiche legate ai principi etici del caos. La forte motivazione che ci ha spinti a scrivere questo libro si fonda proprio su questa idea artistica, che accosta la pratica della ricerca a quella dell'arte e che pone il ricercatore nel ruolo di artista indagatore, metadiffusore dell'informazione; colui che si preoccupa, in maniera laterale e creativa, non tanto di limitarsi a produrre qualcosa in rete, ma di offrire prospettive trasversali che permettano di fruire delle esperienze che già esistono e soprattutto che stanno evolvendo.

Non si tratta di difendere a spada tratta il dominio sulla tecnica come fossimo in un Nuovo Illuminismo Cibernetico, ma dal momento che la tecnica, per certi versi, è destinata al dominio, ci sembra utile e necessario operare in maniera tale che questo dominare sia in mano dei molti e non dei pochi. Non vogliamo che pochi dominino la tecnica per dominare i molti, ma che la tecnica sia dominata da tutti per non essere dominati dalla tecnica.

1.2 MEMORIA E CONOSCENZA

La disorganizzata ed esagerata quantità di informazione presente su Internet, se da un lato offre a tutte le persone collegate la possibilità di fruire di conoscenze e saperi disparati, dall'altro lato può generare linee di condotta disorientanti se non addirittura repulsive. Gli utenti della rete, vedendo questo sconfinato orizzonte informativo, spesso fuggono in preda al panico temendo di potersi in qualche modo smarrire (*loss of orientation*), allontanandosi dalla fonte, dal continuo fluire dei dati informativi, non considerano il fatto che la non gerarchizzazione del modello comunicativo potrebbe viceversa aumentare la possibilità di ricevere nuove informazioni. Da questo, potremmo dedurre una sorta di semplice equazione definita in funzione della quantità di informazione ricevuta: la nostra possibilità di memorizzare informazione risulta inversamente proporzionale alla verticalizzazione della diffusione dell'informazione, mentre è direttamente proporzionale all'orizzontalità del parametro distributivo dell'informazione. Un esempio può essere questo: nel momento in cui conduco una ricerca in rete su uno specifico argomento, la possibilità di avere a disposizione i dieci risultati più importanti elaborati da un dispositivo verticale (come potrebbe essere una redazione umana selettiva e rigida, o un motore di ricerca che recensisce i siti) mi preclude al contempo la possibilità, seppur aleatoria, di trovare altri cento risultati, tra i quali estrarre cose nuove e importanti di cui non verrei mai a conoscenza, tramite un motore di ricerca fondato su funzionamenti di scelta dei risultati più aperti ed elastici. Se vale questo esempio, ne deduciamo che il processo orizzontale comunicativo fondato su una struttura agerarchica (che non smetteremo mai di ricordare è uno dei punti naturali e vitali di Internet) potrebbe essere una condizione favorevole per l'accrescimento della conoscenza sia individuale sia collettiva.

Chiediamoci: che cosa sono memoria e conoscenza in rapporto all'informazione? Che cosa significa informazione giusta e sbagliata, buona e cattiva? Iniziamo con queste due domande per addentrarci in riflessioni successive.

La visione, ormai completamente accettata da tutti, che il computer sia un'estensione e un'esteriorizzazione della memoria induce a riflettere su un duplice aspetto: sulla memoria globale in rapporto ai crescenti e veloci apporti delle scoperte scientifiche, sulla memoria individuale in stretta relazione all'informazione telematica. Il primo aspetto conferma il fatto che oggi giorno la mnemotecnica è ormai in stretta relazione con le invenzioni della scienza moderna; Internet, visto in senso lato, e i motori di ricerca in particolare, sono alcune delle forme più significative e capitali di organizzazione della memoria collettiva contemporanea. Ora, se pensiamo a Internet come figura dominante (dominio del-

la tecnica) e agli individui come figure dominate, slittiamo sul secondo aspetto. In altre parole siamo di fronte al forzato adeguamento dell'esperienza singola a quella collettiva: a una crescente supremazia della tecnica corrisponde una memoria crescente del singolo, e quindi una gestione sempre più funzionale ed efficace della memoria. Intuitivamente capiamo la maggiore centralità del motore di ricerca e la sua ragione di "comodità" per il singolo utente: per non essere inondato dal mare informativo io, navigante, sono garantito dal motore, a scapito della perdita di una certa autonomia di scelta.

Noi concepiamo la memoria come discernimento, immagazzinamento e disfacimento dell'informazione. Non si tratta quindi di considerare la memoria come un semplice elaboratore di carico/scarico dell'informazione, ma come un dispositivo in grado di costituire una rete autonoma interpretativa in grado di confrontarsi e connettersi al "generale".

Il meccanismo di ricerca in rete è quindi mosso da una doppia dinamica. Da una parte, siamo di fronte a dinamiche di sazietà informativa, grazie al materiale vagliato e preconfezionato ricevuto (lavoro svolto dai motori di ricerca), dall'altra non si deve sottovalutare l'esigenza che muove l'utente di sapere la "verità" di ciò che riceve. Ma cercare una definizione di "verità" all'interno del caos telematico porta a dover fare i conti con veri e propri ossimori, quali quello di *verità caotica*. La verità sta nel momento operativo, nella paziente sequenza del processo creativo della ricerca, nella consapevolezza della diversità dell'esperienze: "Perfino il biologo più ostinato, tuttavia, dovrebbe ammettere che il mondo è in molti modi diversi – in realtà ci sono anche molti diversi mondi esperienziali – a seconda della struttura dell'essere interessato e del genere di distinzione che esso è in grado di compiere. E anche se restringiamo la nostra attenzione alla conoscenza umana, si può pensare che il mondo abbia molti modi di esistere. Questa convinzione non oggettivista (e, nel migliore dei casi, anche non soggettivista) si sta lentamente facendo strada nello studio della conoscenza. Finora, tuttavia, questo orientamento alternativo non ha avuto un nome ben definito, in quanto è poco più di un ombrello che offre riparo a un gruppo relativamente ristretto di persone impegnate in campi diversi. Noi proponiamo di usare il termine *enattivo* per dar risalto alla crescente convinzione che la conoscenza non sia la rappresentazione di un mondo prestabilito, ma piuttosto l'enazione, la produzione, di un mondo e di una mente sulla base della storia delle diverse azioni che un essere compie nel mondo. L'approccio enattivo, quindi, prende seriamente in considerazione la critica filosofica all'idea secondo cui la mente sarebbe uno specchio della natura, ma si spinge oltre, rivolgendosi a questo problema dal nucleo stesso della scienza".[3]

La ricerca telematica non riguarda quindi solo una forma astratta di

conoscenza, ma è essa stessa una forma di esperienza e conoscenza, che comprende in sé la forte volontà di scoprire qualcosa di nuovo.

Essere felici di "ciò che passa il convento" o cercare da altre parti?

Legata a questa si pone anche l'altra questione, come distinguere l'informazione buona da quella cattiva? Chi deve decidere rispetto alla qualità dell'informazione? Le decisioni prese singolarmente nel differenziare l'informazione giusta da quella sbagliata sono sufficienti a risolvere il problema per intero? La maggior parte dei partecipanti alla vita del grande network è in grado di rispondere a questa necessità, oppure ha bisogno di trovare già tutto pronto? Occorre che alcuni prendano le decisioni per gli altri? Chi?

Come non ci sono regole e basi precise sull'organizzazione e fruizione della conoscenza in rete, che non ci stancheremo di ripetere è solo finzione di caos, così non esistono sistemi e comportamenti strutturati nell'elaborazione e collocazione logica dell'informazione nel nostro cervello. In entrambe le situazioni, individuale e collettiva, è l'esperienza e l'interpretazione creativa a modificare le reali interconnessioni sia a livello neuronale del singolo cervello, sia in riferimento all'insieme di più cervelli, producendo capacità di acquisizione e spostamenti di saperi. Si tratta di un concetto di esperienza connettiva che prevede dei modelli di reciprocità e correlazione. Tutto ciò che avviene all'interno del singolo (cervello) e nella comunità (Internet) deve essere in rapporto alle attività di tutte le componenti (nuclei, strati, livelli, nodi ecc.) seguendo una logica di cooperazione, in condizione di rendersi autostabilizzante e di controllarsi e proporzionarsi autonomamente.

Forse una chiave di volta del problema relativo alla conoscenza in rete è rappresentata dall'analisi del comportamento dei motori di ricerca, caratterizzati da forme di filtraggio, esclusione, moderazione, "censura".

1.3 ECCESSI E CENSURE

L'accezione "classica" della parola censura rimanda alla volontà esplicita di un soggetto, o più soggetti, generalmente istituzionali, mossi da una intenzionalità punitiva. Se vogliamo, il nostro lavoro, si distacca da quella lettura "classica", per diversamente tentare di porre l'accento sull'attività a tratti decisamente ambigua e "censoria" svolta dai motori di ricerca.

I milioni di utenti che quotidianamente usano Internet non si accorgono di certo, a meno che non si occupino attentamente del problema, degli aggiustamenti impercettibili, quasi "subliminali", messi in atto da chi svolge lo smistamento dall'informazione telematica. È un "sano" metodo di censura, che i diretti responsabili giustificano con la necessità di fare pulizia, di dare autorevolezza ai loro servizi, peraltro ben ac-

colti da molti utenti. È un metodo che attribuisce "verità", riducendo la molteplicità delle risposte possibili a pochi risultati rispetto alla grande vastità esistente in rete, ma che in realtà nasconde e vela ragioni prettamente commerciali e politiche.

In termini più spiccioli bisognerebbe chiederci: perché mi deve essere tolta l'opportunità di scegliere?

Se ci facciamo trasportare nella rete dai vari accompagnatori: browser, motori di ricerca, portali ecc., senza un atteggiamento critico e diffidente, vacillerà il principio cardine, fin dai suoi esordi, della comunicazione telematica rappresentato dalla necessità dell'interazione con gli altri esseri umani pensanti (soprattutto attraverso la produzione, lo scambio e la fruizione delle idee per mezzo delle pagine web).

Nel manipolare e canalizzare la libera, "confusionaria" circolazione del materiale comunicativo, viene inibito il pensiero del singolo. Si tratta di una strategia di controllo digitale, che potrebbe essere attivata dai poteri forti per garantire sicurezze ed equilibri già dati. Il "sistema del potere" non corre mai grossi rischi in realtà, tranne forse in qualche serio attacco da parte di "pirati informatici".

Ma come abbiamo già sottolineato, il fenomeno del "diluvio" informativo pone interrogativi sui gradi di discriminazione in atto del sapere. Si possono elencare tre categorie, o gradi, di sentenze censorie: risoluta, mediata ed invisibile.

Il *controllo risoluto* prevede *interventi duri*, si pensi all'esempio dell'Arabia Saudita, che impedisce agli utenti di collegarsi individualmente a Internet, se non previo il rilascio di permessi e licenze, e *interventi meno duri*, come in Cina e a Singapore, dove le maglie dei server statali bloccando il traffico proibito stilando una sorta di "lista nera" con i nomi dei trasgressori. È per certi versi scontato elencare ragioni e motivazioni di questa censura, visto il tipo di politica in vigore in questi stati, anche se le loro frontiere elettroniche si stanno sempre più aprendo – come nel caso di SINA.com (http://www.sina.com), portale cinese. In Occidente invece la linea dura venne lanciata con il Communication Decency Act – CDA – che la Corte suprema degli Stati Uniti, nel giro di meno di un anno, ha dichiarato incostituzionale in quanto violava la libertà di espressione dei cittadini (giugno 1997), decretando il clamoroso insuccesso del CDA, con relativo effetto boomerang.

In Occidente i sistemi di filtraggio delle informazioni rispondono a una logica più *mediata*. Con questo termine intendiamo identificare in particolare l'attività di quei gruppi e associazioni, più o meno folkloristici, di cybervigilanti che filtrano l'informazione "proteggendo" gli utenti, principalmente dal proselitismo neonazista e dallo sfruttamento sessuale dei bambini, e inoltre intendiamo particolari software di bloccaggio dell'accesso.[6] Un'altra articolazione di questa forma di censura è incentrata sull'ipotesi di costituire dei supervisori dell'informazione,

che giudichino cosa cestinare o cosa distribuire; con la previsione di istituire dei "centri di qualità", "centri di garanzia", dei buffoneschi "comitati di garanti", specializzati in merce digitale, come se si trattasse di bottiglie di olio o vino sulle quali dover apporre l'etichetta DOCG.[7] Un'idea garantista assai opinabile e a tratti folle, che cela la presunzione di giudicare per gli altri, ma soprattutto nasconde la volontà di costruire una vera e propria oligarchia culturale.

L'ultima categoria di filtraggio, o censura, è quella *invisibile* e chiama in causa tutti i motori di ricerca. Tramite gli aggiustamenti del ranking sui risultati ottenuti su una parola chiave, si giudicano e si classificano qualitativamente le pagine web. Gli strumenti per ottenere un ottimo posizionamento (attualmente viene considerato un buon ranking l'essere inclusi nelle prime venti posizioni) sono sia economici sia tecnici. Nel primo caso, si assiste a lauti esborsi a favore dei motori di ricerca, per aumentare o in caso per mantenere, la propria posizione (acquisto di keyword, accordi personali ecc.). Il secondo è esito di conoscenze tecniche, soprattutto di programmazione e impostazione del documento, per aggirare gli agenti software. In entrambe le situazioni l'azione di filtro resta nascosta agli utenti, ma quel che è certo è che molti siti esclusi dai risultati pur essendo di qualità eccellente, non saranno minimamente considerati dall'utente-ricercatore. Una censura silenziosa e anonima, quindi, ma che a pensarci bene produce situazioni di contenimento ed indirizzamento, che avvantaggiano solo certe situazioni culturali.

Gli interessi economici e politici che stanno dietro alle varie forme di "censura" tendono a nascondere inoltre problemi di carattere etico. Tra questi la necessità di tutelare la libertà individuale di produrre, diffondere e fruire informazione digitale. Una libertà continuamente minacciata, e pertanto una questione che va assumendo sempre più l'aspetto di un terreno carsico, dove forze ed elementi sotterranei lavorano per corrodere e distruggere questo bene prezioso. Spetta a ognuno di noi, con grande sacrificio, se non vogliamo essere pilotati come automobiline, occuparsene per mantenere viva questa libertà. Come primo passo, capire e accettare che il concetto di ricerca in rete ha un significato più ampio e variabile certo meno gerarchico di quello tradizionale, e che occorre costruirsi via via le proprie tecniche personali e di autodifesa digitale. Al momento attuale siamo un po' tutti come Sisifo: per una giusta causa, portiamo il nostro macigno fino in cima alla montagna e poi qualcuno lo respinge indietro. La sfida è in atto.

1.4 TECNICHE, SAPERE, INDIVIDUO, LIBERTÀ

L'accessibilità al sapere telematico e allo stesso tempo la visibilità dei website all'interno del WWW sembrano essere gli assi portanti sui quali verte il rapporto tra l'uomo contemporaneo e le sue potenzialità di ac-

quisire e fornire nuove conoscenze in rete. Internet costituisce quel luogo-contenitore aperto dove sono convogliati, depositati e relazionati tra loro gli innumerevoli punti di vista degli individui: idee, nozioni, progetti, desideri, pensieri, teorie...; posizioni individuali che connesse tutte insieme generano un sapere collettivo strutturato orizzontalmente, fruibile in modo autonomo senza costrizioni dirette.

La struttura stessa della rete presenta dei mezzi, e insieme dei metodi tecnico-scientifici, detti "risorse di ricerca": dallo stesso browser di navigazione, passando per i motori di ricerca fino al semplice scambio tra gli utenti dei bookmark personali, che permettono di accedere più o meno liberamente alle informazioni.

Il motore di ricerca è visto in particolare dalla maggior parte degli utenti come un dispositivo sicuro e affidabile, una sorta di oracolo telematico, dal quale ci si aspetta estrema velocità e garanzia nelle risposte. L'utente medio, di fronte all'interfaccia pulita, amichevole e funzionale del motore di ricerca utilizzato, proietta il proprio bisogno di sicurezza e di razionalità scientifica, non preoccupandosi assolutamente di cosa ci sia dietro al semplice form in cui immette la richiesta. L'utente medio, tuttavia, dovrebbe ricordarsi che i principali motori di ricerca sono delle organizzazioni puramente commerciali, delle aziende che trattano la gestione delle informazioni, che a tutt'oggi procurano immensi guadagni e le cui quotazioni in Borsa influenzano pesantemente l'andamento dei maggiori indici internazionali.

Proprio la capacità dei naviganti di conoscere le principali tecniche di uso delle risorse di ricerca, le modalità di selezione e scelta del materiale, i trucchi e le malizie per migliorare il ranking, sempre in continua modificazione, costituiscono un sapere parallelo al sapere diffuso, un sapere non gerarchico, anzi caotico come quello rappresentato da tutta l'informazione esistente su Internet. Il sapere diffuso è la massa di informazioni; il sapere parallelo è costituito dall'insieme di conoscenze che permettono di appropriarsi del sapere diffuso. È possedere la piena conoscenza dei dispositivi, dei criteri e delle procedure di ricerca: il navigare nelle forme più elementari, il riconoscere quando un link è utile oppure inutile, l'individuare i motori di ricerca e i portali più autorevoli, il sapere ricercare sia in maniera semplice sia avanzata, il frequentare newsgroup, mailing-list, forum e chat, il capire cosa è nascosto dietro a un buon posizionamento nei primi risultati. Saper fare tutto ciò significa essere in possesso di un metasapere. Questo metasapere può essere semplicemente definito come l'arte di saper scegliere: in primo luogo di ciò di cui abbiamo bisogno; in secondo luogo, la tecnica di scelta. Un sapere che richiede specifiche competenze, funzioni tecniche e un'apertura mentale che non guardi più al centro come al luogo fondante dal quale si origina il sapere, ma volga lo sguardo

verso un insieme di periferie diramate qua e là, delle quali il soggetto conoscente deve rintracciare i reciproci collegamenti. L'assenza di un centro nella rete deve indurci a trattare i risultati delle ricerche come esperienze non prestabilite, discontinue, che si intrecciano o si sviluppano parallelamente, si ignorano o si separano, ma che in generale tendono a moltiplicarsi vertiginosamente. Occorre inventarsi un approccio personalizzato "a togliere" nei confronti del flusso informativo, che ci permetta di evitare sia l'informazione inquinata sia la troppa informazione. Inoltre, bisogna avere chiara una cosa: il motore di ricerca non ci mostra sempre il suo volto leggibile e sincero, non lavora per disporci le cose a nostro piacimento e favore. Noi dobbiamo immaginare il fare ricerca come un distorcere, un violentare, un dubitare; tale esercizio dovrà dar luogo a una serie combinatoria di eventi di cui noi stessi fisseremo i contorni. Se, in generale, i risultati della ricerca devono essere concepiti innanzi tutto come insieme, come grappoli di risultati, quale legittimazione e statuto dobbiamo attribuire ai risultati da noi scelti?

Bisogna accettare il rischio dell'incertezza come paradigma di base che sostenga la logica della ricerca su Internet. Si tratta di realizzare un dominio sulla tecnica (in questo caso la tecnica è rappresentata dalle risorse di ricerca e dalla conoscenza nel mezzo) per controbilanciare il dominio della tecnica che il mondo dell'economia capitalista usa per manipolare a proprio favore il potere nell'organizzare e distribuire l'informazione.

Lo sforzo individuale che tutti dovremmo compiere consiste quindi nell'apprendere l'arte dello scegliere – per far sì di non essere scelti. In un certo senso il dominio della tecnica deve essere inteso dal soggetto come un pericolo per la propria libertà di pensiero e di azione.

Dove esiste un pericolo deve esserci anche qualcosa che ci faccia scampare da tale pericolo. La grande libertà che Internet ci mette a disposizione deve essere vissuta come un premio; per guadagnarselo è necessario faticare sia individualmente sia collettivamente.

NOTE

1 I primi codici del linguaggio HTML – Hypertext Markup Language – e le regole della trasmissione HTTP – Hypertext Transport Protocol – prendono vita agli inizi del 1990 a opera dell'informatico inglese Tim Berners-Lee in collaborazione con il francese Robert Cailliau (cfr. Tim Berners-Lee, *Il futuro della Rete*, Feltrinelli, Milano 2001).

2 Su questo concetto cfr. Pierre Lévy, *Il diluvio informazionale*, http://www.emsf.rai.it/dati nella sezione "Aforismi"; P. Lévy, *Cybercultura*, Feltrinelli, Milano 1999, p. 156; Ralph Begleiter, *Le déluge d'informations*, Dossiers mondiaux, "Revues électroniques de l'USIA", vol. 1, n. 12, settembre 1996.

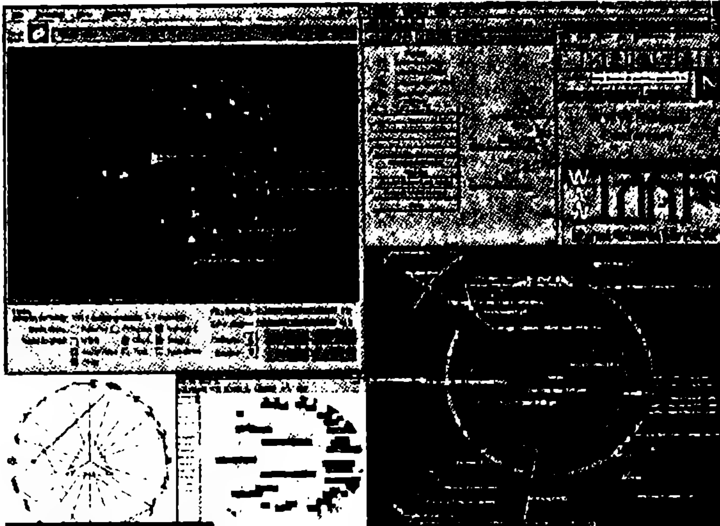
3 Sui processi dell'entropia, probabilità e incertezza applicate al futuro della scienza cfr. Ilya Prigogine, *La fine delle certezze: il tempo, il caos e le leggi della natura*, Bollati Boringhieri, Torino 1997.

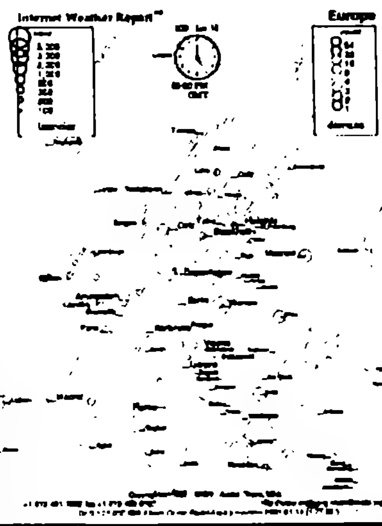
4 Con "scatola nera" o black box, si tocca un punto cruciale della lettura critica data dagli hacker alla tecnologia. Ispirandosi al concetto elaborato dalla scuola cibernetica, con "scatola nera" viene inteso nella letteratura "mainstream" un costrutto tecnologico che deve essere assunto acriticamente, senza possibilità di smontaggio e quindi di interpretazione critica. A questa visione della tecnologia, che è poi quella dominante nella società, si contrappongono le pratiche hacker che proprio nello "smontaggio" hanno il loro perno fondativo.

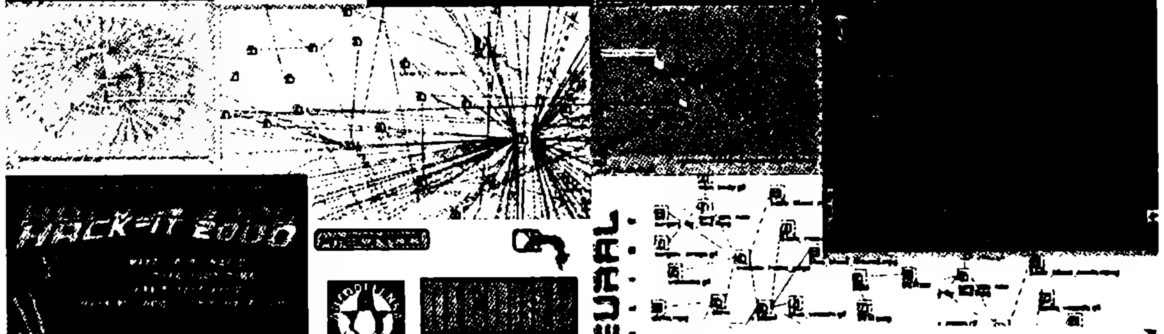
5 F.J. Varela, E. Thompson, R. Eleanor, *La via di mezzo della conoscenza*, Feltrinelli, Milano 1992, pp. 31-32.

6 "I programmi di bloccaggio tradizionale, come SurfWatch, CyberPatrol, NetNanny, CyberSitter, X-Stop e molti altri, funzionano in base a due diversi principi. Un primo sistema, rudimentale e aspramente criticato, vieta l'accesso a tutte quelle risorse digitali individuate per mezzo di una ricerca per parole-chiave, con il rischio documentato, per l'attiguità semantica di molte parole "hot", di privare l'utente di informazioni spesso preziose, come quelle relative alla prevenzione dell'Aids o persino di testi letterari e di riproduzioni digitali di opere d'arte. Un secondo sistema, più evoluto, impedisce di collegarsi con una serie di siti considerati poco adatti per i minori in base a categorie stabilite dalle aziende produttrici di software. Anche questo sistema è oggetto di numerose critiche: primo, perché le scelte vengono effettuate spesso da personale poco qualificato; secondo, perché queste scelte rispondono in ogni caso a criteri del tutto soggettivi; terzo, perché a una impossibile operazione di screening completa si sostituisce un'analisi superficiale e grossolana; quarto e ancora più importante, perché i siti effettivamente bloccati sono spesso di contenuto politico. Per esempio, CyberPatrol blocca l'accesso a numerosi siti di gruppi estremisti e radicali e di associazioni femministe e di omosessuali; X-Stop blocca l'accesso persino ai siti del think tank conservatore Heritage Foundation o a quello della setta religiosa dei Quaccheri"; (da Internet, *La ragnatela della libertà*, "L'INQUISITORE DIGITALE", di Giuseppe Mancini, http://www.ideazione.com/Rivista/1998/maggio_giugno_1998/mancini_3_98.htm).

7 Cfr. l'intervista a Umberto Eco, *Le notizie sono troppe, imparate a decimarle, subito*, sulla rivista "Telèma", n. 4, 1996, http://www.fub.it/telema/TELEMA4/Eco4.html; e la polemica, suscitata dalle posizioni dello stesso Eco, con Roberto Casati nell'articolo *Vade retro, esperto*, "Il Sole 24 ore", 5 marzo 2000. Eco ribadisce la sua critica a Internet anche all'edizione 2000 della fiera informatica milanese Smau, dove con un intervento orale ha chiesto un "sigillo", un "watermark" per i siti validi rispetto a quelli poco buoni.



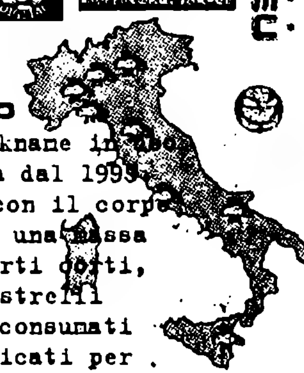






hs

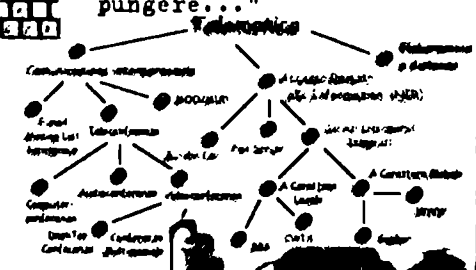
hacaro



Hacaro, nickname in uso
in Italia dal 1999.
"Animali con il corpo
formante una massa
unica, arti corti,
polpastrelli
spesso consumati
e modificati per
pungere..."







Il LOA
hacklab MI
al deposito Bulk

LOA

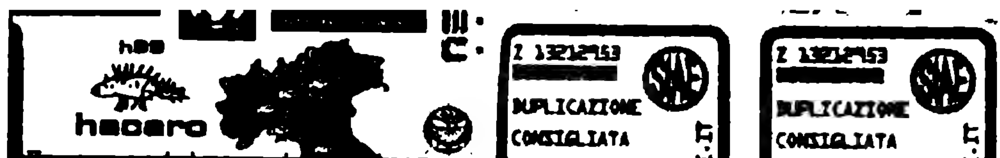
Z 13212953
[redacted]
DUPLICAZIONE
CONSIGLIATA
ANY MEDIA
NO SIAE
BULK PRODUCTIONS
037
LOA Hacklab - MI
SIAE
WWW.NETSTRIKE.IT

Z 13212953
[redacted]
DUPLICAZIONE
CONSIGLIATA
ANY MEDIA
NO SIAE
BULK PRODUCTIONS
037
LOA Hacklab - MI
SIAE
WWW.NETSTRIKE.IT

Z 13212953
[redacted]
DUPLICAZIONE
CONSIGLIATA
ANY MEDIA
NO SIAE
BULK PRODUCTIONS
037
LOA Hacklab - MI
SIAE
WWW.NETSTRIKE.IT

Z 13212953
[redacted]
DUPLICAZIONE
CONSIGLIATA
ANY MEDIA
NO SIAE
BULK PRODUCTIONS
037
LOA Hacklab - MI
SIAE
WWW.NETSTRIKE.IT

UNSIAE 4 bollini
contro il virus del copyright



2

ORGANIZZAZIONE RIZOMATICA DELL'INFORMAZIONE IN RETE

2.1 STRUTTURA DELLA RETE

La rete è come ce la immaginiamo. Né potrebbe essere altrimenti, essendo la rete una serie infinita di oggetti, simboli e interpretazioni, materiali e immateriali, collegati fra loro nei modi più bizzarri e complessi.

Per Marcos[1] "la rete siamo noi tutti che parliamo e ascoltiamo" – ed è sicuramente una definizione su cui riflettere –; per molti utenti storici della telematica la rete è l'insieme delle aree di discussione delle BBS,[2] o al massimo i famigerati newsgroup.[3] Per la maggior parte delle persone la rete è il Web,[4] per altre ancora la posta elettronica. Per chi scrive è un'interconnessione rizomatica, un qualcosa che non si fa in tempo a descrivere che già cambia profondamente, come un sistema biologico in un percorso coevolutivo.[5]

Ma la rete è innanzitutto costituita da un numero enorme di computer di vario tipo che costantemente, oppure occasionalmente, approfittano di un vettore fisico (che può essere il comune cavo telefonico, ottico ecc.) e di un modem[6] per comunicare fra di loro.

Ovviamente, non tutti i computer hanno le stesse funzioni e possibilità comunicative. In Internet, la distinzione principale in una comunicazione di tipo client-server[7] è fra i computer che ospitano i servizi (per esempio quelli che ospitano i domini,[8] e che uno studio dell'Uni-

versità di Berkeley, in California, stima intorno ai 13.000.000 di unità) e i computer che usufruiscono di tali servizi. Questi sono i computer che possono leggere le pagine web appartenenti a tali domini: un numero elevatissimo e in crescita esponenziale, impossibile da determinare!

Una volta fotografata l'attuale situazione tra fornitori di servizi e utenti collegati fra loro fisicamente perlopiù attraverso una ragnatela di cavi, ma anche attraverso l'etere, rimane da sapere come avviene la comunicazione fra macchine. In telematica, gli insiemi di regole che controllano la comunicazione tra macchine vengono definiti protocolli. Il più noto tra i protocolli è il TCP/IP (Trasmission Control Protocol/Internet Protocol) che permette a una vera e propria infinità di macchine di attivare sessioni di comunicazione nei modi più disparati. Le più conosciute e praticate forme di sessioni sono il Web, la posta elettronica e i newsgroup, l'IRC e l'FTP (che rappresentano rispettivamente le modalità storiche di chat e trasferimento di file in rete), così come Telnet (accesso diretto a computer remoti), senza tralasciare le moderne possibilità di attivare sessioni audio e video attraverso nuovi algoritmi di compressione come l'MP3 oppure tramite programmi proprietari come il REAL-AUDIO-VIDEO.

Il protocollo di comunicazione attuale, l'IPv4, permette già delle funzioni mirabolanti che hanno fatto la fortuna di Internet, assicurando la comunicazione in tempo reale fra soggetti molto distanti fra loro. Ma la sua futura implementazione, già annunciata da tempo – l'IPv6 – consentirà un aumento considerevole del numero di macchine che potenzialmente possono accedere alla rete, insieme a un'ottimizzazione e controllo del traffico della rete stessa.[9] Tutto ciò si attua attraverso una traduzione alfabetica del numero di riconoscimento della macchina in rete, detto appunto numero IP, delegato a una rete di server speciali noti come DNS (Domain Name Server), e a un'organizzazione gerarchica della rete. In tal modo, una macchina accede alla rete solo attraverso un'altra macchina già presente in rete e che ha già acquisito il credito e la configurazione necessaria per essere visibile al resto della rete. Viene così a configurarsi un network, sotto un profilo logico, tutt'altro che orizzontale e anarchico ma piuttosto piramidale e verticistico.

2.2 COMUNICAZIONE, OVVERO COME METTERE IN COMUNE L'INFORMAZIONE

Se per accedere alla rete bisogna passare attraverso un sistema di accreditamento basato su una gerarchia pluripiramidale i cui vertici sono praticamente irrintracciabili e, peraltro, difficilmente gestibili, data la complessità della struttura della rete, è pur vero che un governo della

rete esiste, anche se ha delle caratteristiche molto particolari, per certi versi, di tipo comunitario.

Le regole tecniche di governo[10] della rete si sviluppano attraverso documenti di proposta detti RCF (Request For Comment) che sono pubblicamente valutati, discussi e, nel caso, modificati da chiunque si presti al gioco di testare le nuove ipotesi di governo tecnico e di sottoporle a sua volta al pubblico esame.[11] Viene quindi a formarsi un network gerarchico fondato su una logica "democratica".

Per capire come si comunica[12] in rete, ovvero come si mette in comune un'informazione, è importante indagare sulle regole basilari che governano la trasmissione bidirezionale di dati che si svolge attraverso la ragnatela dei cavi, ma anche tramite il sistema ipertestuale dei link.[13]

Entrando in rete sempre e comunque come ospiti di macchine gerarchicamente più importanti o più accreditate della nostra, ereditiamo da queste la nostra legittima capacità di banda, cioè la velocità di connessione (quantità di dati trasmessi per unità di tempo), che dovrà bastarci per effettuare – in tempi funzionali ai nostri scopi – comunicazioni di varia natura: dalla navigazione sul Web alla posta elettronica, fino alle videoconferenze. Ogni sessione di comunicazione ha una sua particolare esigenza di banda; quindi, mentre per utilizzare Telnet, o la posta elettronica, i newsgroup e IRC sono sufficienti poche risorse, per navigare efficacemente sul Web dobbiamo disporre di una buona velocità di connessione; se poi vogliamo scaricare file tramite FTP, ascoltare un brano musicale o partecipare a una sessione video, necessitiamo di una banda sempre più estesa.

Operando in Internet, non è tuttavia necessario conoscere solo le nostre risorse di banda, ma anche quelle dei nostri interlocutori: essendo la rete un ambito di scambio comunicativo bidirezionale, pur avendo eventualmente la fortuna di disporre di una veloce linea dedicata, dovremo comunque preoccuparci delle possibilità tecniche del nostro interlocutore. La reciprocità di condizioni è irrinunciabile: se entrambi gli interlocutori non hanno la "banda" necessaria per attivare l'auspicata videoconferenza, forse dovranno più realisticamente ripiegare nel più semplice ma sempre intrigante dialogo in IRC...

La condivisione dei saperi richiede una "reciprocità tecnologica" nell'accedere al medium.

Cosa c'entra tuttavia questo preambolo con l'operatività dei motori di ricerca?

La struttura della rete, le sue caratteristiche tecniche e le modalità di interconnessione delle macchine che la costituiscono determinano capacità operative e limiti dei motori di ricerca, essendo anch'essi, sotto forma di collettivi redazionali o di spider automatici, degli eterni e insaziabili browser in cerca di informazione in giro per il Web. I motori di ricerca risentono infatti non tanto di (improbabili) loro limiti di con-

nettività, ma delle scarsa larghezza di banda – nota anche come "bandwidth" – della fonte di informazione da indicizzare, tanto che un sito risulta più difficilmente indicizzabile dai motori quando è ospitato su un server con poca banda.

2.3 I MOTORI DI RICERCA E LA RETE

I motori di ricerca sono macchine complesse basate sia su risorse umane sia su automatismi. Essi hanno il compito di scandagliare una porzione prescelta della rete, indicizzarla e metterla a disposizione di ipotetici utenti. Le prestazioni e la geografia della rete incidono senza ombra di dubbio sulla funzionalità dei motori di ricerca.

Se la velocità di connessione fra macchine comunicanti in rete ha un'importanza notevole nell'economia della comunicazione in Internet, è la geografia dei link a giocare un ruolo di primo piano nel tentare di dare forma e senso alla rete.

Nel 1999 Albert Réka, Hawoong Jeong e Albert-Lazlò Barabasi pubblicano su "Nature"[14] uno studio secondo il quale la distanza tra due qualsiasi "pagine web" è mediamente compresa in un percorso di soli 20 link. Come al solito le statistiche però nascondono una situazione molto complessa.

Uno degli studi più esaurienti e aggiornati sulla topografia della rete è stato svolto dall'IBM Almaden Research Center in collaborazione con il Systems Research Center e la società Altavista di San Matteo (California) e presentato alla IX International World Wide Web Conference ~ (http://www.almaden.ibm.com/cs/k53/www9.final) svoltasi ad Amsterdam nel maggio 2000.

Secondo questo studio, la rete si presenta come un corpo centrale di 56 milioni di pagine web strettamente interconnesse, ovvero pagine che sono molto linkate le une con le altre, seppure secondo criteri e percorsi diversi e di differente lunghezza. Da questo corpo centrale si staccano due protuberanze che assieme raccolgono altri 44 milioni di pagine: la prima è costituita da pagine personali e siti minori, i quali sono linkati al corpo centrale – portali, motori di ricerca, siti d'informazione e pagine delle aziende più rinomate – ma che non godono del link di ritorno; la seconda è costituita invece dalla maggioranza dei siti delle imprese, delle università, dei centri di ricerca e di altre istituzioni che beneficiano di link dal corpo centrale, ma non contraccambiano con link di ritorno. Collegamenti di leggera entità connettono poi le due protuberanze minori, mentre sono numerose le pagine completamente isolate dalla rete.

Come incide tutto ciò sull'operatività dei motori di ricerca? Moltissimo, ma soprattutto per due aspetti fondamentali: il reperimento dell'informazione e il ranking.

Tanto per cominciare un'informazione, per essere presa in considerazione da un motore di ricerca, deve essere prima reperita. Nel mare magnum valutato genericamente in centinaia di milioni di pagine web, reperire le informazioni diventa un compito continuo estremamente complesso. Per esempio, c'è molta differenza fra un'informazione contenuta in una pagina non linkata e una che è identificata tramite link, da molte altre pagine e che facilmente, prima o poi, sarà intercettata dal laborioso navigare dello spider in questo oceano digitale. Le pagine linkate infatti hanno il pregio di essere considerate dai motori pagine più interessanti e aggiornate della rete, al contrario di quelle isolate, la cui unica speranza di essere identificate dal motore di ricerca resta quella di farsi notare o per un caso fortuito o grazie a una segnalazione extra ipertestuale che si spera sia presa nella dovuta considerazione.

Tra questi due estremi – molti link, nessun link – c'è una varietà di modi di inserire link nelle pagine e organizzare le pagine in siti con vari livelli di profondità,[15] che incide, per le ragioni ora espresse, sulla possibilità o meno per una pagina web di essere trovata da un motore di ricerca.

Un altro importante aspetto relativo all'indicizzazione è rappresentato dal criterio utilizzato dai motori di ricerca per decidere come stilare la classifica dei risultati conseguiti dopo un'interrogazione: il ranking. È considerato un buon risultato (top ranking) quello con alta visibilità per l'utente in cerca di informazioni. Apparire almeno fra i primi venti risultati di una ricerca effettuata su un motore è ormai l'aspirazione di ogni sito in cerca di visibilità.

Evitiamo per il momento di considerare quei posizionamenti – di solito i primi posti negli indici creati dai motori – ottenuti tramiti abili strategie di marketing. Tali strategie, pur se ipotizzabili,[16] sono di fatto non facilmente praticabili e risultano comunque molto costose. Resta il fatto che tanto maggiore è la quantità di link presenti in un sito quanto migliore è la reperibilità e visibilità del sito in rete. E ciò, è chiaro, favorisce la possibilità di un buon posizionamento entro le venti posizioni negli indici dei motori, sempre più ambite, sia da grandi sia da piccole aziende.

Nell'attuale geografia di Internet, un'informazione contenuta in una pagina web ha buone probabilità di essere reperita da un motore di ricerca solo se è posizionata in un server che ha buone prestazioni in termini di larghezza di banda. Ma soprattutto l'informazione sarà tanto più al centro dell'attenzione dei motori quanto maggiore sarà il numero dei link, altre pagine web, che rimandano a quella pagina. Cosicché non solo quella pagina risulterà facilmente reperibile dagli spider, ma essendo considerata "popolare" nel microcosmo delle consultazioni digitali,[17] otterrà buone probabilità di essere inserita nel top ranking dei motori di ricerca.

2.4 SI FA PRESTO A DIRE INFORMAZIONE...

L'informazione in rete, malleabile e volatile come un bit digitale, può avere caratteristiche fra le più differenti. La pagina web, costruita utilizzando un "linguaggio a marcatura" fra quelli considerati accettabili dalla rete (per esempio il comunissimo HTML), padroneggia fra i possibili file presenti in rete, e spesso viene identificata con la rete stessa. Tuttavia, in Internet sono numerosissimi gli oggetti presenti: pagine testo in formato .txt; miriadi di immagini in formato .gif, .jpg o magari .png; database dai formati più diversi; programmi software; video e audio, i cui formati sono costantemente in evoluzione.

Il cercare di risolvere in vari modi il problema incontrato dai motori di ricerca nel catalogare informazione elaborata nei formati più diversi, è un problema immenso.

Per esempio creando appositi servizi: Altavista ha una sezione specifica http://image.altavista.com per la ricerca delle immagini; mentre Lycos ha una sezione http://mp3.lycos.com, utilizzata per trovare i file audio di maggiore successo. Esistono quindi motori specializzati nella ricerca di oggetti specifici.[18]

Dal punto di vista del reperimento dell'informazione, invece, oltre ad adottare risorse umane di tipo redazionale e utilizzare automatismi,[19] alcuni motori di ricerca si avvalgono di algoritmi di analisi dell'immagine in grado di riconoscere segnali, testo e didascalie nascoste nel codice.[20] Si tratta di informazioni normalmente usate per la sincronizzazione e la descrizione, per esempio, nelle trasmissioni satellitari, oppure per finalità di titolazione e ricerca interna a opera di software specializzati.[21]

Resta il problema di catalogare informazioni di tipo testuale (file di testo), HTML[22] non più e non solo in termini qualitativi, ma anche in termini quantitativi. Accenniamo di nuovo al problema di localizzare l'informazione nella rete, ma anche a quello di acquisire e catalogare razionalmente ed efficacemente l'informazione dopo averla localizzata. Il primo aspetto è stato "risolto" saltando alla ovvia conclusione che è praticamente impossibile localizzare tutta l'informazione presente sulla rete; si è quindi in un certo senso deciso di lasciare emergere dalla trama magmatica e ridondante della rete un'insieme di classifiche compilate in bell'ordine nello stile pop di una hit parade. Anche la soluzione della seconda problematica risulta tuttavia improbabile e non offre garanzie.

Al di là della bontà e dell'efficacia degli algoritmi di indicizzazione usati dai motori, rimane infatti irrisolta la questione per cui quasi tutti gli spider[23] dei motori di ricerca acquisiscono solo la parte iniziale dei siti che visitano, senza scandagliarne la profondità e i contenuti reali.

Possiamo quindi affermare, trist* e sconsolat* che i motori di ricerca

che credevamo infallibili ed esaustivi sono in realtà dei laboriosi ma grossolani procacciatori di informazioni, incapaci di districare – senza censure, insabbiature o macchinosi artifici – la trama caotica e ridondante della rete.

NOTE

1 Dalla *Seconda Dichiarazione della Realidad per l'umanità e contro il neoliberismo,* Una rete intercontinentale di resistenza per l'umanità, Dalle Montagne del Sudest messicano, Comitato Clandestino Rivoluzionario Indigeno Comando Generale dell'EZLN La Realidad, Pianeta Terra, agosto 1996.

2 BBS sta per Bulletin Board Systems e sono sistemi di comunicazione telematici amatoriali con caratteristiche tecniche particolari e non omogenee fra di loro. Ormai molto rari, perlopiù si riconoscono per la loro attitudine libertaria, localistica e per essere in rete fra loro non in tempo reale.

3 Si sa: quello che non si conosce fa paura, ed è per questo forse che i newsgroup hanno una così brutta fama (tra le persone che ne ignorano la logica). Simili alle aree di discussione delle BBS, i newsgroup rappresentano invece, per molte persone esperte della rete, la matrice, ovvero la parte viva di Internet, il flusso incontrollato di informazioni e dibattiti dove misurare lo stato di salute della rete.

4 Il Web, o World Wide Web (W3), è la rappresentazione multimediale della rete attraverso il protocollo HTTP. È lo strumento più immediato per rapportarsi a Internet.

5 Tommaso Tozzi, *Cotropia, lifeware e coevoluzione mutuale*, La Stanza Rossa, Bologna 1997.

6 Modem o modulatore-demodulatore è uno dei tanti apparecchi di uso più o meno comune che permettono ai nostri pc di collegarsi a Internet attraverso un processo di trasformazione del segnale che permetta di attraversare il vettore scelto.

7 Per comunicazione client-server si intende un particolare modo di comunicazione nel quale vi sono dei soggetti che forniscono dei servizi, e altri, i fruitori, che usufruiscono dei servizi. Tale tipo di scambio di informazioni avviene anche durante la navigazione di pagine web, quando il nostro browser (come Communicator di Netscape, Explorer, Opera ecc.) è il fruitore (client) e il computer che ospita le pagine il fornitore (che ha quindi installato un server web adatto allo scopo).

8 Il suffisso di primo livello che appare all'estremità destra di un nominativo Internet e che qualifica la nazionalità (per es. it per Italia) o il tipo di servizio (per esempio org per organizzazione no-profit) di provenienza.

9 L'IPv6 proprio per via delle sue alte prestazioni ha suscitato qualche perplessità fra gli amanti della privacy, in quanto utilizzando l'IPv6 sono ipotizzabili anche delle procedure che identificano facilmente gli utenti della comunicazione.

10 Per un approfondimento sul funzionamento e sull'organizzazione della rete sia in termini tecnici sia burocratici cfr. anche Luciano Paccagnella, *La comunicazione al computer*, il Mulino, Bologna 2000.

11 Malgrado queste forme di autogoverno e il senso di libertà che è avvertibile appena si accede alla rete, molte persone e collettivi si sono preoccupati, nel corso degli anni, di ipotizzare o sviluppare reti autonome che avessero meno strumenti di controllo di quelli presenti in Internet. Ciò anche per meglio sfuggire a quei casi di censura dell'informazione come quello rappresentato nel 1998 dalla momentanea chiusura del server di movimento di Isole nella Rete ~ www.ecn.org, che ha

fatto molto discutere della rete e delle sue potenzialità (anche) repressive. Per esempio, è opinione comune che prendendo il controllo dei principali nodi di routing della rete possono essere intercettate praticamente tutte le informazioni della rete! "Il progetto GAIA ~ www.ecn.org/loa/terminal/GAIA è nato per risolvere un problema frequente nelle reti a organizzazione gerarchica: la dipendenza da chi detiene un livello di accesso privilegiato alla struttura piramidale". Paul Garrin, un media-artist di New York, ha progettato un network alternativo di DNS in modo che cambiando il settaggio della propria macchina sulle coordinate di uno dei server di Garrin si possa cambiare le regole alfabetiche della rete. È di Tommaso Tozzi la proposta di un'INTERFACCIA FLUTTUANTE che raccolga al suo interno l'indirizzo di tutti i server nel mondo senza nessun criterio di scelta o censura e con la possibilità per chiunque di aggiungervi l'indirizzo del proprio server. La proposta più forte rimane comunque quella denominata Freenet ~ http://freenet.sourceforge.net/ Freenet è costituita da una rete di macchine che pur "appoggiandosi" a Internet adottano un proprio protocollo di comunicazione che permette loro di ricercare, immettere e scambiare informazione all'interno di questa rete, senza che ci siano possibilità tecniche di censura o di rintracciamento dell'autore dell'informazione. Ovviamente ciò è possibile facendo uso di quello che è indubbiamente una delle risorse più interessanti della telematica attuale ovvero della crittografia cosiddetta "pesante".

12 L'etimo di "comunicare" secondo alcuni filologi deriva dal latino: cum (insieme) e munus (dono). La comunicazione forse come scambio paritario e condivisione di saperi.

13 I link sono quelle porzioni di testo dei documenti in rete su cui è possibile clickare con il mouse per passare in tempo (quasi) reale da una zona di un documento a un'altra, oppure da un documento a un altro, indipendentemente dalla loro localizzazione geografica.

14 Fa ancora più scalpore la notizia, riportata sempre su "Nature", a proposito della scarsa capacità dei motori di ricerca di indicizzare la rete. Il record di indicizzazione, circa il 14% dell'intera rete, sembrerebbe spettare a Northern Light: www.northernlight.com. E il restante 86%?!?

15 La profondità di un sito è determinata dalla quantità di differenti livelli di link che esso contiene.

16 Fa parte della nostra proposta artistica il provare ad aggirare alcuni limiti dei motori di ricerca. L'approfondimento di questo argomento verrà riportato nell'ultimo capitolo di questo libro.

17 L'indice di "popolarità" (ovvero il numero di pagine che linkano al sito preso in esame) è uno dei parametri più importanti adottati dai motori di ricerca per considerare di qualche interesse l'indicizzazione di un sito. Un altro parametro, anch'esso oramai molto conosciuto, è costituito dall'aggiornamento della pagina: più una pagina viene rinnovata in grafica, ma soprattutto in contenuto, e più il motore è invogliato a mantenerla nel proprio elenco di siti da visitare, visitandola con una frequenza proporzionale alla frequenza di rinnovo del sito registrato.

18 TUCOWS ~ www.tucows.com è la famosa directory per risorse software. Ancora FTP SEARCH ~ ftpsearch.ntnu.no/ftpsearch interagisce con molti archivi file in FTP e I-SURF ~ http://isurf.interpix.com è specializzato nel recupero di immagini.

19 È un aspetto sul quale torneremo più volte; i risultati prodotti dai motori di ricerca sono il frutto di molteplici attività e competenze. La ricerca automatica effettuata da software dedicati è spesso affiancata da un lavoro redazionale condotto da équipe di professionisti.

20 Un esempio di Audio e VideoLogger di nuova generazione è costituito dalla tecnologia sviluppata da www.virage.com e utilizzata, per esempio, anche da Altavista.

21 Nel caso di un file MP3 viene estrapolata l'informazione normalmente usata dal let-

tore hardware MP3 oppure dal software di lettura presente sul computer per dare un titolo al brano e verificarne autore e lunghezza.

22 Mentre per un'informazione di tipo testuale è possibile un'indicizzazione che fa riferimento a liste di vocaboli conosciuti, per un file di tipo binario risulta molto più difficile decodificare l'informazione che l'autore ha voluto trasmettere, magari sotto forma di prestazione di un software o sequenze di immagini (vedi capitolo sull'accessibilità dell'informazione).

23 Spider è un programma del motore di ricerca che viene lanciato nella rete alla ricerca di informazione: lo si può immaginare come un browser molto semplice e veloce incaricato di navigare e registrare grosse quantità di pagine web secondo rotte preordinate dal programmatore che lo ha lanciato.



T. Berners-Lee



H. Bey



I. Prigogine



G. Boole



G. Salton



I. Clarke



L. Torvalds



J.P. Barlow



L. Zadeh



R. Penrose



P. Feyerabend



L. Felsenstein



R. Stallman



L. Wittgenstein



G. Deleuze

3

CENTRALITÀ INEDITA DEL SAPERE

3.1 OLTRE IL SAPERE COME CONTENUTO

Il sapere: concetto che accompagna da sempre le vicende umane, che ha assunto, e assume nel tempo posizioni diverse, se è rapportato alle evoluzioni scientifiche, sociali e politiche. Il sapere muta secondo le modalità di comunicazione che lo contengono, le svolte epocali delle società, le decisioni politiche dei governi. Per leggere la complicata storia delle forme del sapere occorre adoperare più filtri e lenti, per non incorrere in interpretazioni deterministiche ed errate, che possono causare forti abbagli se non addirittura accecamenti. La parola sapere ha due significati costitutivi fondamentali che la contraddistinguono.

Un significato del sapere è di natura propriamente contenutistica; un sapere-contenuto che rappresenta tutto il bagaglio di informazioni e conoscenze esistente in relazione a un periodo storico preciso, a un'area geografica, a un popolo e alle sue tradizioni, e ai mezzi di comunicazione usati per la trasmissione; si pensi alla differenza sostanziale che c'è tra la diffusione del sapere tramite il libro oppure via etere. L'altro significato riguarda un aspetto strutturale e metodologico delle tecniche usate per distribuire l'informazione; descrive l'insieme del primo tipo di sapere; ovvero, esso designa l'insieme di conoscenze relative all'uso degli strumenti per accedere e acquisire sapere, cioè il metodo –

come direbbe Descartes: "Per metodo intendo regole certe e facili... per aumentare il sapere e pervenire alla vera definizione delle cose". Questo sapere-metodo ha la caratteristica di essere variabile e necessita, per il proprio sviluppo, di continui aggiornamenti. Questi due ruoli assunti dal sapere: produzione e contenuto, circolazione e accesso, anche se ben distinti spesso coincidono, si coordinano e sono collegati tra loro nelle proprie funzioni.

Nell'ambito della comunicazione telematica le due forme di sapere appaiono piuttosto distinte. Da una parte abbiamo l'insieme delle informazioni che sono freneticamente prodotte per la rete, il sapere-contenuto; dall'altra abbiamo l'insieme di metodologie che rendono possibile la diffusione, la classificazione e il reperimento di tali informazioni, il sapere-metodo. Quest'ultimo tipo di sapere ha un ruolo rilevante, se non determinante, nella comunicazione tra computer. La dimensione tecnica diventa primaria, coloro che detengono le conoscenze specifiche al riguardo producono una sorta di monopolio del sapere tecnico che fortunatamente è abbattuto o quantomeno ridistribuito grazie alle caratteristiche del mezzo tecnologico usato, la rete. Stiamo parlando di un sapere che difficilmente si pone il problema di essere conservato, viste le sue continue modificazioni. Infatti, proprio a causa della sua immediatezza e fluidità, la comunicazione telematica si presta poco alla conservazione.

3.2 UN SAPERE CENTRALE E MOBILE

Essere capaci di rendersi visibili in Internet, cioè conoscere bene le regole che ci garantiscano un ottimo posizionamento nei risultati dei motori di ricerca e sapersi districare nella caotica fruizione del materiale divulgato dalla rete, tutto ciò costituisce di per sé un sapere. Come abbiamo già detto, un sapere-metodo che permette a chi frequenta il mondo telematico di far fronte al momento ("nei prossimi cinque minuti")[1] ai due problemi nodali: la visibilità e la ricerca in rete.

Al sapere enciclopedico,[2] decentrato e "prelevabile" da chiunque scorre parallelo il sapere inteso come puro know-how, conoscenza tecnica detenuta da poche persone: "Il mouse è la nuova chiave a stella, è lo strumento polivalente per conoscere e possedere il mondo. Nella rete concreta del territorio, c'è un sapere che la scuola di per sé non possiede".[3] Il sapere assume una posizione centrale ed è mantenuto in una roccaforte nella quale non tutti possono entrare.[4] Espressa così brutalmente quest'idea potrebbe apparire una gratuita provocazione, una difesa di coloro che sono contrari a qualsiasi innovazione scientifica. Invece crediamo che questa possa rappresentare una spinta positiva per migliorare la nostra vita all'interno della rete. Ne è prova evidente il sempre maggiore aumento delle pratiche comunicative che si basano

sulla filosofia del *work in progress* e sulle *risorse free*,[5] e che connotano soprattutto il sapere pratico, il sapere come strumento. Il know-how ha stretta necessità di essere condiviso per migliorarsi continuamente, e non ha nessun bisogno di essere custodito gelosamente come fosse un mistero esoterico.[6] L'eccezionalità del sapere individuale viene meno, a favore di un processo di cooperazione nell'acquisizione e nella costruzione di un sapere di gruppo, che diventa un patrimonio collettivo aperto e pronto a ricevere costanti apporti, sia demolendo sia ampliando il già conosciuto.

Il know-how caratterizzato da differenze e specializzazioni assume dunque un ruolo centrale e decisivo nell'aspetto situazionale della comunicazione, dell'arte e soprattutto delle imprese. La centralità del sapere si situa al contempo in un'area allo stesso tempo mobile, provvisoria e flessibile. Sarebbe impossibile pensare alla conoscenza informatica come a un qualcosa di definitivo, di durevole per tutta la vita: sovente bisogna azzerare quello che abbiamo appreso per ripartire di nuovo.[7] Questa flessibilità influisce profondamente nei rapporti interni al mondo del lavoro; le strutture rigide e gerarchiche di produzione sono costantemente scalfite dalle individualità dei singoli i quali si sentono sempre più responsabili e coautori di un unico processo. Non si deve credere però che sia oro tutto ciò che luccica; gli orizzonti della New Economy non fanno sparire con un colpo di bacchetta magica le disuguaglianze e lo sfruttamento del lavoro. L'astrattezza di certi investimenti e operazioni commerciali destinate a sicuro fallimento ricadono sempre sui più deboli, sui singoli lavoratori, costretti a camminare sempre più sul filo del precariato, dell'insicurezza e del rischio. Gli slogan che molti imprenditori della rete, a più riprese ripetono spesso: "la rete siete voi", "Internet è di tutti", eccetera, non riescono concretamente a far sparire i fantasmi peggiori del fordismo.[8]

3.3 IL SAPERE COME COLLANTE SOCIALE

Questo libro esamina la visibilità e la ricerca in rete, che insieme rappresentano una buona parte di tutto quel sapere tecnico-pratico indispensabile per la comunicazione telematica. Questo know-how è caratterizzato da tre aspetti peculiari. Esso è cumulativo, cooperativo e flessibile. Il canale maestro dove è trasmesso il know-how è Internet: website, forum, newsgroup su temi specifici; che insieme ad altri luoghi: spazi associativi, ambiti di formazione nel mondo del lavoro, meeting[9] ecc., stabiliscono dei rapporti sociali che si fondano su relazioni reciproche di scambio delle conoscenze. Un sapere incarnato nella condivisione delle competenze, che è l'esatto contrario del sapere riconoscibile solo nel modello istituzionale: autorità statali, scuole, monopoli editoriali, concorsi. L'istituzione assoluta che abbisogna costantemente di autorevo-

lezza, che si preoccupa di difendere le posizioni forti ed elitarie del sapere, contribuendo a sostenere il suo potere culturale per paura di perderlo. La detenzione e la salvaguardia del sapere diventano manifestazioni di potenza a scapito delle relazioni sociali, il più possibile paritarie tra gli individui. Il sapere è legittimato e conservato dall'istituzione la quale si riproduce per mezzo del sapere stesso.

Il know-how di cui abbiamo parlato, si rivela, invece, un componente propulsivo per lo spazio di relazioni sociali all'interno delle comunità, l'elemento principale che permette la comunicazione di rete che implica la collaborazione tra più persone. È una forma di cooperazione del sapere dove ognuno riceve secondo i suoi bisogni e offre secondo le sue capacità. È la messa in connessione delle differenze individuali disposte a essere collettivizzate, senza che esse perdano la propria originalità e autonomia. Questo valore relazionale del sapere alimenta la vita sociale in Internet a esso collegata, secondo il paradigma tipico del networking, cioè quello capace di mantenere in vita una comunità attraverso lo scambio da pari a pari per fine condivisi. In una comunità così costruita c'è una propensione a mettere in comune le conoscenze, a pensare il network come luogo di decentramento del soggetto singolo, di dissoluzione delle identità partecipanti, dove il proprio io è "un io flessibile, un collage di frammenti sottoposti a un incessante divenire, sempre aperto a nuove esperienze";[10] il lavoro di costruzione del sapere viene inteso alla stregua di "si è fatto qualcosa", e non come un "ho fatto qualcosa". Questo valore di rete è una vera e propria fonte di ricchezza, frutto delle "utopie" del mondo cyberpunk, alla quale oggi si abbeverano anche le imprese commerciali della nuova economia. Una ricchezza sociale che va sempre più trasformandosi in strategia imprenditoriale di guadagno.[11]

3.4 CHI NON PARTECIPA NON SA, CHI NON SA NON PUÒ PARTECIPARE

Prima del saper fare è necessario il voler saper fare, e questa necessità si manifesta con lo sforzo di acquisire strumenti, cioè, in pratica, con la partecipazione alle vicende del sapere in rete. Il mondo di Internet comprende milioni di persone che si scambiano informazioni, sentimenti, sapere ed emozioni; per questo la nostra parte di vita digitale deve essere presa sul serio. La comunicazione elettronica ha come scopo principale il mantenimento e l'evoluzione delle relazioni tra le persone, connesse in rete, riconosciute attraverso dei codici (c.parrini@leonet.it, www.strano.net) trasmessi con diversi dispositivi (e-mail, pubblicazioni web e altro).[12] Atti relazionali che avvengono nonostante l'estrema riduzione di fisicità propria della "normale" convivenza sociale: quanti abbracci, odori, strette di mano verranno meno, e quante opere d'arte oggettuali vanno scomparendo.

Coloro che si troveranno al di fuori delle comunità elettroniche saranno come pesci fuori dall'acqua, piante senza luce. Molti antropologi, sociologi e filosofi contemporanei sostengono pessimisticamente che la futura vita digitale porterà con sé solitudine e depressione, non tenendo conto del fatto che sarà proprio l'assenza di uso di nuove tecnologie dell'informazione a isolare gli individui dalla comunicazione, dalla conoscenza, fonti indispensabili per l'esistenza, con altrettanti prevedibili risultati catastrofici. Stabilendo relazioni di particolare interesse con nuove persone, inserendosi all'interno di comunità dedicate, la nostra percezione del reale andrà incontro a drastici mutamenti; non solo accrescerà, in generale, la varietà dei punti di vista su determinati temi, ma ci permetterà anche di estendere il nostro punto di vista, di dichiarare la nostra posizione, in qualche modo di esprimerci. Può accadere che una certa comunità non accolga, non "aspetti" un individuo-utente, oppure che lo stesso soggetto abbia l'esigenza di gettare le basi per aprire una nuova comunità (magari con diverse logiche di comportamento) dove convogliare altri individui con interessi analoghi, innescando un meccanismo di mutualità a favore di uno spirito di convivenza più aperta, contro ogni forma di standardizzazione.

Si tratta comunque di necessità strettamente legate alla questione del "sapere tecnico". La possibilità di trovarsi esclusi da questo sapere è quindi in stretta relazione con la nostra volontà di partecipare collettivamente non tanto alla produzione di risposte astratte, quanto alla definizione di regole di sopravvivenza.

NOTE

1 La frase è lo slogan preso in prestito da *Next five minutes*, conferenza e festival sull'arte, politica, attivismo e media tattici. La conferenza N5M nasce nel 1993; la terza si è tenuta ad Amsterdam e Rotterdam, tra il 12 e il 14 marzo 1999; http://www.n5m.org.

2 *L'Encyclopédie* segna lo sconfinamento di una lettura universalistica del sapere, cara alla cultura seicentesca, verso un'impostazione classificatoria delle conoscenze, registrando i legami fra le varie discipline, assumendo così la funzione di un "mappamondo" del sapere.

3 V. Foa e A. Ranieri, *Il tempo del sapere*, Einaudi, Torino 2000, p. 42.

4 Anche linguisticamente il sapere tecnico-informatico è scollato dalla comunicazione linguistica in generale. Talvolta, parlando con programmatori o sistemisti di rete, ci troviamo di fronte a forme di vero e proprio bilinguismo: i termini e le frasi usati, infatti, sono tutt'altro che elementari e comuni.

5 Cfr. l'esperienza Linux. Creato da Linus Torvalds nel 1991, Linux è un sistema operativo *free-resource*, in cui i sorgenti di programmazione sono aperti agli interventi di chiunque voglia fornire integrazioni per migliorarne le funzionalità. Cfr. http://www.linux.org. È da rilevare che nel 1999 la giuria del Prix Ars Electronica di Linz ha assegnato proprio al software *Linux* la Golden Nica, con la motivazione: "Si fonda sui contributi di migliaia di programmatori che volontariamente collaborano

online in uno sforzo di gruppo che ha creato un sistema operativo rimarchevolmente robusto", citata in "Internet News", n. 11, anno V, p. 53. Cfr. anche *Gli alberi di conoscenze*, dispositivo di rete concepito da Michel Authier e Pierre Lévy (Feltrinelli, Milano 2000) che si pone l'obiettivo di sviluppare la conoscenza all'interno di comunità virtuali, lavorando sul concetto di intelligenza collettiva, riconoscimento di saperi, apprendimenti cooperativi. "Poggiando sulle autodescrizioni degli individui, un albero di conoscenze rende visibile la molteplicità organizzata nelle competenze disponibili in una comunità" (P. Lévy, *Cybercultura*, Feltrinelli, Milano 1999, p. 176).

6 Non è un caso che una forma *ante litteram* di procedura operativa *free-resource*, circa la costituzione e trasmissione del sapere, è rappresentata dal sapere medico nel V sec. a.C. Negli scritti egizi e babilonesi veniva usato uno stile manualistico, che seguendo una concezione cumulativa del sapere, la condivisione delle analisi personali dei medici e la formulazione di quadri clinici, giungeva ad allargare il sapere già conservato, aggiungendo la descrizione di nuove malattie, diagnosi e terapie riportate nei primi cataloghi-manuali. In questo caso, il sapere medico è inteso come sapere tecnico, rispetto ad altri saperi di carattere teorico quali il sapere giuridico e quello filosofico, che continuavano a mantenere una comunicazione orale. Per trovare qualcosa di scritto bisogna giungere ai romani (Bruto, metà del II sec. a.C.).

7 "Due considerazioni nei confronti di chi opera nella rete. 1) Guai a sentirsi superiori, nel cybermondo si riparte sempre da zero; 2) Guai a trascurare il cliente, nella nuova economia ha ancor più ragione che nella vecchia". Intervista a Michel Duponchel, direttore dello sviluppo IBM di "@_city", Costa Azzurra, "Corriere della Sera", 27 marzo 2000.

8 In buona sostanza il campanello d'allarme suona quando "all'idea che le reti aperte siano più disponibili a reinvenzioni radicali rispetto a gerarchie piramidali come quelle che dominavano l'era fordista", si affianchino le pratiche del *reengineering*, "reingegnerizzazione" o "ristrutturazione", con il loro aspetto più evidente che è il *downsizing*, cioè la riduzione dei posti di lavoro (R. Sennet, *L'uomo flessibile*, Feltrinelli, Milano 2000, pp. 47 e *passim*).

9 Cfr. l'Hackmeeting, incontro annuale autogestito dedicato agli aspetti sociali, politici e creativi della telematica. Sono previsti incontri, seminari e conferenze su: crittografia, anonimato, sistemi *open-source*, controllo, sperimentazione, hackeraggio sociale, telematica low tech ecc.; si realizzano eventi artistici con video, performance, concerti e dibattiti. Il primo Hackmeeting si è tenuto il 5/6/7 giugno 1998 presso il CPA FI-SUD a Firenze, il secondo il 18.19.20 giugno 1999 al LSOA Deposito BULK di Milano, il terzo il 16/17/18 giugno 2000 al CSOA Forte Prenestino di Roma. Vedi http://www.hackmeeting.org.

10 Richard Sennet, *L'uomo flessibile*, cit., p. 134.

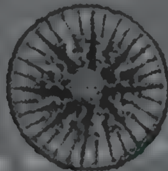
11 Come sottolinea Carlo Formenti citando Kevin Kelly: "...il messaggio di Kelly ai manager dell'informatica suona più o meno così: imitando gli hacker che volevano regalare computer e programmi al popolo, costruirete il vostro successo, perché 'il trionfo del popolo' sarà il *vostro* trionfo. Chi non capisce questa verità e crede di poter campare dei successi passati, anche se è ancora al vertice, ignora di non essere mai stato così vicino alla rovina. Rovina che colpirà tanto più in fretta chiunque si sforza di costruire 'sistemi di proprietà' chiusi, per proteggere prodotti e know-how dalle imitazioni della concorrenza..." (Carlo Formenti, *Incantati dalla rete*, Raffaello Cortina Editore, Milano 2000, p. 216).

12 Agli inizi del 2000 il professore americano Norman Nie attraverso lo Stanford Institute for the Quantitative Study of Society ha realizzato un dossier-sondaggio usando come riferimento a un campione di 4113 soggetti di età che superano i 18 anni: ne viene fuori che oltre il 55% di essi ha accesso a Internet, di questi solo il 15% trascorre meno di un'ora la settimana online, il 49% da 1 a 5 ore, il 22% da 5 a 10 ore, il restante 14% supera le dieci ore settimanali di collegamento.

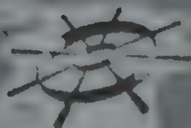
PARTE SECONDA

RICERCA DELL'INFORMAZIONE IN RETE

I motori di ricerca,
per la loro natura fisiologica
e struttura meccanica,
rappresentano il dispositivo
primario per accedere
all'informazione







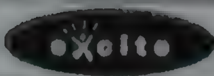


Nel marzo 2000 gli
utenti di AltaVista
hanno raggiunto la cifra
di 54 milioni, effettuando
26.000 ricerche al minuto,
corrispondenti a 37 milioni
di ricerche al giorno.

MotoriDiRicerca.IT













4

IL MOTORE DI RICERCA

4.1 RICERCA DELL'INFORMAZIONE IN RETE. COMUNITÀ, AGENTI INTELLIGENTI, BOOKMARK E MOTORI

La ricerca di informazioni in rete è espletata utilizzando molteplici strumenti che sono comunemente definiti motori di ricerca. In realtà quando cerchiamo in rete un'informazione, forse senza rendercene conto,[1] utilizziamo una serie di strumenti dotati di caratteristiche e funzioni molto diverse fra loro, e tra di essi solo pochi possono essere propriamente classificati nella categoria dei "motori di ricerca".

Secondo la nostra interpretazione, i principali strumenti di ricerca possono essere suddivisi in quattro classi distinte: le comunità in rete; i bookmark o directory; gli agenti intelligenti o strumenti push; e infine i veri e propri motori di ricerca.

Per *comunità* in rete si intendono i newsgroup,[2] le chat,[3] le mailing-list,[4] e tutto ciò che si usa in genere per interloquire con altri soggetti umani. Spesso è in questi ambiti che si trovano le notizie più fresche e aggiornate, o le riflessioni più originali rispetto a argomenti di pubblico dibattito. Ciò che contraddistingue la comunità da altri ambiti più chiusi e gerarchici, è un generoso quanto impressionante spirito collaborativo – improntato alla solidarietà e alla cooperazione orizzontale – che storicamente ha origine negli atteggiamenti e nelle atmosfere di

aperto sperimentalismo maturate in seno alla scena degli anni Settanta delle prime comunità hacker che coltivavano e attuavano il (bi)sogno di praticare una reale socializzazione dei saperi.[5] Oggi troviamo un esempio di questo spirito di copartecipazione in quelle particolari situazioni in cui qualcuno in difficoltà lancia una richiesta di aiuto – perlopiù appelli di natura tecnica per risolvere problemi di funzionamento hardware o software – e subito riceve risposte esaurienti da persone altamente qualificate. Le comunità – come pratica di comunicazione orizzontale –, nella maggior parte dei casi, rappresentano anche un tool, o meglio, un contesto in cui reperire informazioni *ad hoc* o risolvere esigenze specifiche.

La classica modalità per reperire informazioni rimane comunque quella di improntare una ricerca sulla base di cataloghi già esistenti, la cosiddetta "collezione di siti", o bookmark, frutto delle memorizzazioni di precedenti navigazioni. Oppure si può procedere facendo riferimento ad analoghi servizi online di tipo pubblico: le directory dove sono raccolti, divisi per categorie di interesse, numerosi siti specializzati in settori anche molto specifici[6] (dalla letteratura cyberpunk all'immersione subacquea, dalle sostanze psicoattive all'esoterismo). Rientrano nella categoria delle directory sia portali (oggi in costante crescita) sia alcuni famosi servizi, sempre di directory, ma pubblicizzati come motori di ricerca: l'italiano Virgilio o l'internazionale Yahoo.[7] Pur integrando nei loro servizi anche la possibilità di reperire dati nello stile e nelle modalità dei motori di ricerca classici, in realtà forniscono principalmente una raccolta di siti suddivisi per categorie, catalogata ed etichettata da apposite redazioni in tutto e per tutto simili alle eclettiche redazioni delle agenzie di informazione proprie degli altri media e, in alcuni casi, composte da centinaia di navigatori professionisti.

Alcuni cybernaviganti particolarmente smaliziati preferiscono invece ricorrere a tecnologie più sofisticate e impegnative: le tecnologie push o gli agenti intelligenti. Questi sono strumenti che una volta opportunamente configurati permettono di cercare per conto proprio l'informazione dispersa nel Web partendo da una rotta iniziale di navigazione (anch'essa tracciata dall'utente e in maniera particolarmente mirata oppure elaborata personalizzando rotte già note agli stessi strumenti di ricerca). A questo proposito i casi più conosciuti ed eclatanti sono quelli relativi ai gettonati servizi informativi di carattere finanziario, notiziari politici e sportivi che provvedono a far arrivare alla scrivania multimediale e telematica di accreditati giornalisti e businessmen, bollettini quotidiani sull'andamento delle borse o anche solo del tempo meteorologico.[8]

Ma allora in che cosa consiste il vero e proprio motore di ricerca, l'oggetto ancora misterioso dal fascino macchinico dal nome ingegneristico di questo libro...?

4.2 FUNZIONAMENTO DEL MOTORE DI RICERCA

Il motore di ricerca è un insieme di strumenti hardware e software funzionali alla ricerca dell'informazione attraverso un procedimento complesso composto da più fasi.

La prima fase è l'individuazione dei documenti su cui effettuare la ricerca ed è svolta dai cosiddetti "programmi robot" o spider. Questa è un'attività molto impegnativa dal punto di vista delle risorse di connettività e di spazio disco tanto più che l'informazione in rete è spesso destinata a mutare contenuto o indirizzo oppure addirittura a scomparire. Questo forse è proprio uno dei lati deboli dei motori di ricerca, poiché i limiti di spazio disco e di connettività obbligano i motori a dover prendere continuamente in esame solo una parte della rete, ignorandone la geografia globale. L'individuazione dei documenti è fatta a partire da una URL iniziale, navigando di link in link[9] e, connettività permettendo, in un lasso di tempo accettabile. I documenti così trovati saranno periodicamente visitati dallo spider, in un ciclo che si aggira sulle 3-4 settimane.[10] Il limite, pur fisiologico, della porzione di rete monitorata, quantificabile, nelle migliori delle ipotesi, nel 10-15% del totale presunto, è una delle ragioni per la quale i motori si prestano all'accusa di controllo o addirittura censura dei contenuti della rete. In base a quali considerazioni si procede all'esclusione di un certo tipo di materiale? Considerazioni culturali oppure interessi di marketing?

Va pur detto, a onor del vero, che, escludendo alcuni motori di ricerca famosi come Virgilio e Yahoo, che in realtà non sono altro che directory create da un lavoro redazionale e le cui strutture e contenuti sono il risultato di scelte umane, la maggior parte dei motori classici fornisce la possibilità di aggiungere un sito qualsiasi all'elenco delle liste da reperire e indicizzare (con comandi del tipo ADD URL/ADD SITE). In altre parole, tali comandi conferiscono al motore un meccanismo di relativa apertura all'utenza finale.

Una volta reperiti i documenti si passa alla fase successiva: l'indicizzazione. L'indicizzazione è svolta in maniera funzionale, attraverso complessi algoritmi che tengono conto della rilevanza e della frequenza dell'informazione contenuta nel documento da catalogare. L'obiettivo finale dell'indicizzazione consiste nel creare un database, una raccolta-catalogo che dovrà essere interrogabile dal navigatore in cerca di informazioni attraverso interfacce user-friendly. Quello che, al di là di questi processi sotterranei, si manifesta all'utente finale è quindi la lista dei documenti stimati più rilevanti rispetto all'interrogazione effettuata, secondo il criterio del ranking.[11]

4.3 PANORAMICA DEI MOTORI ESISTENTI

Senza voler essere provinciali, iniziamo con due esempi "italiani": Virgilio,[12] "il bello di Internet", un motore di ricerca in versione directory fra i più pubblicizzati e conosciuti; e Arianna, storicamente "il motore di ricerca italiano". Il primo, essendo estremamente selettivo nel suo catalogo scelto da una redazione, è anche molto scarno di informazione. Viceversa Arianna, essendo un motore di ricerca nel senso classico appena descritto, copre una grande fetta di rete, oltre 5 milioni di pagine web indicizzate, e ha la capacità di soddisfare oltre cinquecentomila interrogazioni al giorno in maniera efficace.

Fra i più particolari in campo "internazionale" è Excite, che a tutt'oggi possiede versioni opportunamente disegnate sulle caratteristiche informative e linguistiche di vari paesi – Australia, Canada, Cina, Francia, Germania, Italia, Giappone, Paesi Bassi, Spagna, Svezia e Regno Unito.

La directory per eccellenza in campo mondiale è senza dubbio Yahoo, sebbene anche il meno noto NorthernLight proponga fra le sue varie versioni di presentazione dei risultati anche un approccio che mostra i risultati nello stile organizzato ad albero della directory, impiegando cartelle (SEARCH FOLDERS) suddivise in 200mila soggetti tematici.

La posizione di un sito nelle liste con i risultati di ricerca (il tanto ambito top ranking) viene misurato prendendo in considerazione la popolarità del sito stesso, ovvero il numero di volte in cui quel sito viene citato e linkato da altri siti in Internet. Questo è un criterio sicuramente efficace ma che nasconde anche un risvolto perverso: il fatto di premiare sempre e comunque i siti più noti, finendo per escludere – e non a torto si può dire censurare – le piccole realtà emergenti e poco pubblicizzate[13] che, secondo questa logica a spirale, appaiono destinate nella maggior parte dei casi a rimanere nell'ombra. Per facilitare la ricerca tra una ormai esponenziale mole di informazioni, i motori di ricerca si sforzano allora di escogitare nuove tecniche. Infoseek, per esempio, si inventò la possibilità di ricercare documenti che siano il più possibile simili a uno in particolare, mentre Direct Hit pone in evidenza i siti più visitati su un determinato argomento.

Google, uno degli ultimi arrivati ma che ha subito conquistato una grande visibilità in rete grazie alla sua grafica semplice e scarna e quasi priva di banner pubblicitari, con l'opzione "I'm feeling lucky" fornisce la possibilità di trovare immediatamente in modo giocoso quanto casuale informazioni su un argomento di interesse. Il successo di tale ricerca è tanto illusorio quanto fortuito, come illustra lo stesso motto di Google. Pure, la strategia di ricerca di Google non solo ha fatto la for-

tuna di questo giovane motore di ricerca ma anche scuola, rispetto agli altri, in fatto di ranking.

HotBot mette a disposizione dell'utente una miriade di variabili di ricerca, tra le quali le più note permettono di ricercare documenti reperiti e indicizzati rispettivamente nell'ultimo mese, nelle ultime due settimane o nell'ultima settimana, sviscerando – come dice la traduzione stessa del nome – la possibilità di trovare materiale "caldo", le notizie più fresche e aggiornate.

Altavista, sia nella versione italiana che inglese, è uno tra i pochi motori che presenta la possibilità di utilizzare comandi specifici di ricerca avanzata come illustrato in alcuni esempi:

"Corvo Rosso": cerca esattamente la stringa "Corvo Rosso" e nulla al di fuori di questa espressione
domain:it: cerca pagine che fanno riferimento solo al dominio .it
link:http://www.shake.it: verifica quanti link secondo il database di Altavista puntano a tale sito
host:www.ecn.org: cerca pagine che fanno riferimento all'host (nodo o server di rete) www.ecn.org
image:che: cerca pagine che contengono presumibilmente immagini del che.

Altavista propone altri tool aggiuntivi, come la possibilità di operare una ricerca mediante l'impiego degli operatori booleiani.[14] Si tratta in parole povere di operatori utilizzati in logica matematica per allargare o restringere un campo di definizione. Per esempio:

corvo AND rosso: cerca "corvo" e "rosso"
corvo AND NOT rosso: cerca "corvo" ma non "rosso"
corvo NEAR rosso: cerca "corvo" e "rosso" dove i termini si presentino non solo nella stessa pagina come per AND, ma anche vicini l'uno all'altro.

Di grande utilità sono infine i metamotori di ricerca,[15] sistemi più semplici dei motori di ricerca, poiché si limitano a ritrasmettere l'interrogazione a decine o addirittura centinaia di motori di ricerca. Nel guadagnare in numero di risposte, si perde tuttavia in qualità; c'è il rischio di effettuare un tipo di interrogazione meno raffinata, dal momento che attraverso i metamotori risulta troppo complesso proporre un'interrogazione che sia sofisticata come quella illustrata per Altavista e, contemporaneamente, anche sintatticamente comprensibile da motori aventi regole di interrogazione anche molto diverse tra loro. Così nei metamotori la ricerca avanzata appare oggettivamente drasticamente ridimensionata.

In una panoramica sui motori di ricerca, appare evidente che molti di questi condividano lo stesso database dal quale attingere le informazioni da presentare nella lista dei risultati di ricerca finale, pur ordinandoli in maniera diversa. A questo proposito una risorsa condivisa da molti motori di ricerca è infatti il famosissimo Inkitomi (www.inkitomi.com).

4.4 COME SI RICERCA L'INFORMAZIONE IN RETE

I motori di ricerca, per loro natura, rappresentano uno strumento di interfaccia tra l'utente, cybernavigante esperto o di primo pelo, e l'informazione presente in rete. Spogliata dalla possibilità della ricerca, la rete si presenterebbe infatti come un immenso oceano informazionale, un campo a maglie larghe la cui struttura è in continuo e fluttuante movimento. Come reperirvi l'informazione? Quali pratiche sperimentarvi? Bisogna intraprendere una serie di operazioni critiche per reperire informazioni in situazioni così complesse, innanzitutto cercando di ovviare alle problematiche e ai limiti finora considerati, nella speranza di usufruire di tutte le opzioni offerte dalla rete e dell'inventiva dei naviganti stessi.

Una buona ricerca in rete deve essere performativa, deve così fondarsi su un criterio efficientistico, nel tentativo di ottimizzare le capacità tecniche del mezzo; ma deve anche essere in grado di integrare l'impiego della tecnologia con la creatività.

Suggeriamo in questo senso un possibile percorso di ricerca:

1) Scegliere con attenzione i termini chiave della ricerca;

2) ipotizzare un dominio www.info_da_cercare.it ma anche i suoi possibili equivalenti .com, .org, .net ecc.;

3) non vergognarsi di chiedere consiglio e aiuto alla comunità di cui si fa parte in rete, sia essa una lista, un newsgroup o un canale IRC; è altrettanto efficace provare a cercare in altre comunità in rete che si presume siano a conoscenza delle informazioni cercate;

4) trovare tutti gli strumenti push in materia, noti per poter ricercare il tipo di informazione auspicata;

5) sfogliare in rete le directory, i cataloghi e i portali pertinenti in materia;[16]

6) interrogare tutti i motori di ricerca e i metamotori conosciuti tramite interrogazioni semplici, ma anche in modalità avanzate (e dove possibile ricorrere alla logica booleana);

7) avvalersi di help online come, per esempio, l'italiano www.vorreisapere.it, sul quale si può inviare una richiesta a un operatore in

carne e ossa che effettua la ricerca per nostro conto, inviandoci poi il risultato per e-mail.[17]

Solo l'esercizio continuo della critica e dell'interazione comunicativa, nell'ambito di una pratica collettiva dei saperi, possono forse esorcizzare i rischi di tecnocontrollo o censura digitale, due possibilità insite nello stesso perverso e cattivo impiego dei motori di ricerca in rete. Nella prospettiva di una proliferazione dell'informazione sul Web che tende all'infinito e della sempre maggiore volatilità dell'informazione stessa, diventa necessaria la pratica della strategia del giocatore che conosce le regole del gioco, ma è anche in grado di stravolgerne il senso perché ne capisce la logica profonda.

NOTE

1 Con questo non vogliamo accusare di dabbenaggine la media delle persone naviganti in rete, ma è pur vero – purtroppo – che si sta affermando un modello per così dire televisivo dell'utilizzo della rete, per cui si svolge automaticamente una serie di funzioni immediatamente utili, senza riflettere sulla natura e sulle caratteristiche di funzionamento degli strumenti di rete, senza quindi tenere conto che tutto ciò influisce negativamente sulla qualità dei risultati ottenuti.

2 I newsgroup sono forum pubblici di discussione accessibili tramite un newsreader, come il famoso FreeAgent che si collega a un newsserver pubblico o, più comunemente, al newsserver del proprio provider di accesso, per visionare i forum di eventuale interesse ai quali abbonarsi, per poter partecipare alla discussione con decine, a volte migliaia di altri utenti. I newsgroup si trovano in tutte le lingue e sui più disparati argomenti. Fra quelli storici e a volte più interessanti è doveroso segnalare quelli contrassegnati con la gerarchia *alt*(ernative).

3 Le chat sono tutti quei momenti in cui si realizza una comunicazione in tempo reale con uno o più interlocutori umani perlopiù attraverso servizi web appositamente creati dallo scopo, ma anche attraverso software molto veloci ed efficaci come l'IRC – Internet Relay Chat – oppure attraenti e flessibili, per le loro caratteristiche multimediali come ICQ – I Seek You. Da notare che gli acronimi che identificano software o servizi in rete si devono sciogliere con la pronuncia inglese, ottenendo così la descrizione del loro funzionamento.

4 Per iscriversi a una mailing-list, ovvero a un dibattito che si svolge attraverso lo scambio di e-mail fra indirizzi di posta personali, è necessario inviare una stringa di comando, che varia da mailing-list a mailing-list, a un programma apposito come Majordomo o Listserver, oppure interfacciarsi ad appositi servizi web che facilitano questo compito di iscrizione.

5 È durante gli anni Sessanta-Settanta, prima della privatizzazione delle scoperte dell'industria informatica di base, che si elabora una pratica collettiva della ricerca di avanguardia nel settore del software e dell'hardware. Ne è un emblema la figura di Felseinstein, eretico pioniere dell'informatica (Raf Valvola Scelsi, *No Copyright*, ShaKe Edizioni, Milano 1994).

6 Il catalogo, secondo l'interpretazione di Bruno Snell (*La civiltà greca e le origini del pensiero europeo*, Einaudi, Torino 1963), è uno dei tratti di matrice comune del pensiero occidentale. In particolare fa la sua prima apparizione nella letteratura arcaica dove prima Omero e poi Esiodo lo adoperano nei loro scritti (celebre il "Catalogo delle navi" nel secondo capitolo dell'*Iliade* e l'elenco delle Muse ne *I lavori e i giorni*) utilizzandolo come elemento teorico di conoscenza, ma anche come elemento esortativo ed espediente narrativo: l'esigenza di creare un ordine (kosmos) sia mentale sia estetico dal caos dell'universo esistente si manifesta nella scrittura del catalogo, che appunto è una delle modalità di ordinare gli oggetti nella forma di lista.

7 È curioso utilizzare termini come nazionale e internazionale in un libro come questo che, trattando di rete, non dovrebbe tenere conto delle frontiere politico-burocratiche: in realtà si intende per "italiano" un servizio realizzato in lingua italiana o in Italia; "internazionale" è invece un servizio che tratta materiale informativo originario di più paesi e in più lingue.

8 Spesso e volentieri si pensa a questi agenti intelligenti nei termini di strumenti molto raffinati ma di difficile gestione. In effetti in alcuni casi lo sono, ma esistono applicativi relativamente semplici da utilizzare come Teleport o Wget che, configurati in maniera specializzata e puntati su siti di partenza efficaci, sono in grado di realizzare report quotidiani di grande efficacia su ogni argomento.

9 Una delle caratteristiche principali del Web è appunto quella di poter effettuare una navigazione ipertestuale che utilizzando i link per passare da un documento a un altro, indipendentemente dalla loro localizzazione geografica o di rete.

10 La periodicità in questione dipende da vari fattori ed è direttamente proporzionale alla frequenza con cui viene aggiornato il sito su cui si trova la documentazione reperita dal motore stesso.

11 Approfondimenti su ogni singolo aspetto di questo paragrafo seguiranno nel corso del libro. Un articolo molto sintetico e chiaro è quello di Giuseppe Attardi, *Ricerca di informazione su Web*, in AIIA Notizie, vol. 12, 1999, n. 2, pp. 24-27.

12 Per visualizzare uno qualsiasi dei motori di ricerca citati in questo paragrafo è sufficiente anteporre www. e posporre .com oppure .it (nel caso degli esempi italiani) al nome indicato, escamotage linguistico che peraltro mette in risalto il processo di massificazione e commercializzazione di questi strumenti.

13 Si potrebbe dire poco linkate in questo caso: ovvero il link come sinonimo di popolarità.

14 Da George Boole (1815-1864) scienziato che inventò quella che poi venne chiamata algebra booleana: la base matematica sulla quale si fonda l'algebra di commutazione dell'elettronica digitale che consente di elaborare i valori "falso e vero" che si riferiscono a proprietà, proposizioni, enunciati, eventi ecc.

15 Un esempio per tutti www.metamonster.com che nella sua versione completa è in grado di interrogare contemporaneamente più di 500 motori di ricerca.

16 È chiaro che il percorso di ricerca può essere differente da quello proposto e può anch'esso essere di natura ipertestuale; per esempio, si potrebbe chiedere nell'ambito della mailing-list a cui si partecipa se esistono directory utili, oppure si può cercare con un motore di ricerca se esiste un canale push utile allo scopo.

17 Anche se attualmente non è ritenuto molto affidabile, questo genere di servizi è emergente e si conquisterà probabilmente in un prossimo futuro una certa notorietà e quindi una fetta di mercato ragguardevole. Negli Stati Uniti, per esempio, sono già numerosi e molto frequentati i servizi educational, di training e di formazione universitaria online che possono, in un certo senso, anche essere considerati complessi sistemi di ricerca-reperimento e acquisizione dell'informazione. Non c'è da stupirsi della (apparente) gratuità di alcuni di questi servizi; in realtà, ogni cosa ha un prez-

zo. È molto probabile che il nostro indirizzo di posta elettronica (insieme ad altri dati significativi, quali provenienza geografica, livello di istruzione, possibilità economiche, abitudini di navigazione, gusti, interessi e hobby) sia associato a un preciso profilo appartenente a una specifica classe di consumo. Non è impossibile dunque che a partire da quel momento il nostro e-mail sia bombardato da numerosi quanto spiacevoli messaggi promozionali.

Guerrilla Marketing

Io sostengo il progetto DISABILI.COM

```
!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
  "http://www.w3.org/TR/REC-html401/loose.dtd"
html lang="it"
head
  meta http-equiv="Content-Type" content="text/html; charset=iso-8859-1"
  meta name="Author" content="Arclele,FerryByte,Samuel"
  meta name="KeyWords"
 content="portale accessibilità, portale disabili, w3c standards, web design
  meta name="Description" content="accessibilita' della informazione su web
  title  Accessibilita' delle risorse web - xs2web  /title
/head

body bgcolor="#000066" text="#ffff00" link="#ffffff" alink="#bbbbbb"
vlink="#ff9900"
!-- http://www.ecn.org/xs2web --

p   /p
hr width="50%"

center
hl   b  Accessibilità delle risorse web  /b   /hl
/center

div align="justify"
p  Il Portale Access To Web (xs2web) nasce dalla necessità di rendere
fruibile, "accessibile", l'argomento accessibilità delle risorse web a una
vasta area di soggetti al fine di facilitare quella necessaria formazione di
cultura e coscienza indispensabile per accettare l'idea che oggi il web può
ancora essere ciò che tutti noi sperimentiamo ogni giorno: una comunità di
individui che intrattengono rapporti sociali. Fra questi soggetti ci sono
persone di tutte le età, cultura, condizione sociale e disabilità.
L'abbattimento delle barriere al libero accesso alle informazioni Web è oggi
più che mai un obiettivo perseguibile.  /p

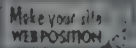
p   /p

center
big  &lt;   a href="news.htm"  News  /a   |   a href="risorse.htm"  Risorse
href="contrib.htm"  Contributi  /a   |   a href="infosite.htm"  InfoSite  /a
href="mappa.htm"  SiteMap  /a   &gt;  /big   /center
/div

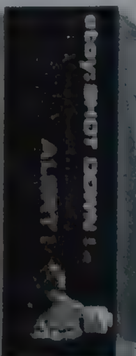
center
br
Ultimo aggiornamento 17 gennaio 2001  /center
br width="50%"
/body
/html
```



A.S.P.H.I.



AS.CI.TEL.



5

IL MOTORE DI RICERCA E L'INFORMAZIONE

5.1 NECESSITÀ DI INDICIZZAZIONE

> "...le carte coprivano già la metà del pavimento. – Si è fatto molto lavoro – disse il sindaco crollando il capo – e quella non è che una piccola frazione. Il più l'ho riposto nel granaio, e veramente la massima parte è andata perduta. Come si fa a conservare tutto? Ma nel granaio c'è ancora molta roba. Potrai trovare il decreto? – disse rivolgendosi alla moglie – Devi cercare un documento in cui la parola agrimensore è sottolineata con la matita azzurra – ..."

Così Kafka ne *Il Castello* (come altri grandi scrittori di un passato ancora del tutto analogico)[1] anticipava non solo le odierne paure di non riuscire ad accumulare e catalogare "tutta" l'informazione, ma addirittura alludeva, con l'accenno alla "sottolineatura azzurra", alla necessità di indicizzarla.

L'indicizzazione diventa una necessità quando, terminata la fase di recupero dell'informazione, bisogna catalogarla e classificarla in modo tale da poterla successivamente interrogare con facilità.

La prima scelta[2] importante che il motore di ricerca effettua avviene sulla porzione di documento da indicizzare: Altavista ed HotBot, per esempio, indicizzano solo i primi 2-300 caratteri del documento, mentre Arianna indicizza il testo nella sua interezza.

Non ci stancheremo mai di far notare che, in maniera diversa ma articolata e comunque utilizzando procedure molto rigorose, tutti i motori di ricerca provvedono a ri-produrre – e allo stesso tempo a deformare – l'informazione. E se il diritto alla deformazione dell'informazione,[3] così come alla sua produzione, è un principio stabilito dalle prime comunità hacker, nelle quali anche noi ci riconosciamo,[4] è altrettanto vero che tale deformazione (o produzione) dovrebbe avvenire secondo procedure (ri)conoscibili dall'utenza finale, e comunque basate su meccanismi che non siano facilmente manipolabili dalle esigenze di mercato o di controllo/censura dell'informazione.

Possiamo ora provare ad addentrarci nella parte più specificamente tecnica del processo di indicizzazione di un motore di ricerca.[5]

La fase di indicizzazione consiste nella creazione di indici cosiddetti *invertiti*, allo scopo di facilitare l'interrogazione. Ma chi indicizza, utilizza le stesse parole di chi effettua l'interrogazione di ricerca sul motore? È un problema dai molteplici risvolti, dal momento che l'interfaccia utente che svolge la ricerca va a interrogare un database a sua volta indicizzato secondo regole precise, che potrebbero anche aver escluso, o comunque non sottolineato, una parola inclusa in un documento, a causa della sua frequenza o collocazione nel documento stesso.[6] Il risultato della ricerca, e soprattutto il criterio di classificazione (ranking) dei risultati con cui sarà fornita risposta all'interrogazione, dipenderà quindi da queste regole.

La fase di indicizzazione corrisponde a un processo di analisi dei documenti che andranno a costituire il database. È in questa fase che le parole portatrici di funzioni sintattiche, e che però non contribuiscono direttamente a specificare un certo contenuto informativo – le cosiddette stopword –, dovranno essere eliminate per poter identificare le parole riconosciute come significative.[7] Dopodiché si giunge alla creazione degli indici invertiti: l'elenco delle parole che assieme costituiscono la porzione di testo in cui si trova la parola significativa, affiancate da un coefficiente che esprime la loro frequenza nel testo.

Attualmente, la maggior parte dei motori di ricerca utilizza l'algoritmo di Porter insieme ad altri algoritmi che sono in grado di ponderare, oltre alla frequenza, anche variabili qualitative come la posizione delle parole citate all'interno del testo.

Ricapitolando, il motore di ricerca seleziona una porzione di rete da cui recuperare l'informazione di ipotetico interesse, e dopodiché analizza questa informazione in toto oppure solo in parte (per esempio i capoversi iniziali).

L'analisi dell'informazione e la sua catalogazione – l'opera di vera e propria indicizzazione – avviene secondo regole formulate in base ad algoritmi matematici, che cercano di ignorare le parole considerate non significative. L'esclusione dei termini avviene filtrando il testo in

questione attraverso una "griglia linguistica", costruita utilizzando appositi dizionari che ignorano tutte le preposizioni, congiunzioni ecc. oppure sulla base di criteri quantitativi, che cioè si riferiscono alla frequenza delle parole nel testo, troppo alta o troppo bassa. Ciò che rimane delle parole considerate significative viene a questo punto "pesato" a seconda della sua frequenza e collocazione nel testo analizzato, ed è in questa fase che localizzazione e reiterazione dei termini giocano un ruolo essenziale.

Facciamo due esempi pratici. Una parola posta all'interno dei tag <h1>[8] è sicuramente interpretata come molto importante, tale quindi da spiccare agli "occhi" digitali dello spider. Lo stesso accade per una parola che appare raramente in una pagina, e molto di frequente in un'altra, anche se entrambe fanno parte della stessa documentazione: lo spider infatti suppone di trovarsi di fronte, rispettivamente, a un indice e a un approfondimento dell'indice stesso, e quindi considera il termine di particolare interesse.

Tutto ciò si traduce nel fatto che il motore, una volta formulata un'interrogazione, produrrà un elenco di risposte stilando una classifica: una lista che cercherà di soddisfare le esigenze dell'utente che effettua l'interrogazione presentando i documenti contenenti la parola chiave richiesta in una graduatoria costruita sulla base di una rilevanza presunta. In altre parole, la presenza della parola chiave in posizioni di particolare risalto, non solo nei metatag ma anche nel testo stesso, e la sua frequenza pregiudica il suo ranking.

Quanto osservato mostra la necessità di produrre un'informazione che sia accessibile al motore di ricerca. Un "tipo di informazione" che possa essere facilmente recuperato e correttamente indicizzato, poiché la visibilità dell'informazione agli occhi degli spider coincide, giocoforza, con l'accessibilità ai motori di ricerca.[9]

5.2 FORME ACCESSIBILI DI INFORMAZIONE

L'informazione deve essere in primo luogo reperibile, quindi interpretabile. Una considerazione quasi lapalissiana che tuttavia non viene presa nelle dovute considerazioni dalla maggior parte dei produttori e distributori dell'informazione in rete.

La preoccupazione che invece accomuna tutti gli autori e i webmaster che bramano diffondere informazione in rete, è quella di confezionare nella maniera graficamente più attraente possibile un oggetto da collocare sull'ambita ragnatela ipertestuale, senza perdere troppo tempo in inutili riflessioni generali.

La smania di poter apparire sulla vetrina più luccicante e fantasmagorica che l'uomo abbia mai potuto immaginare fa spesso dimenticare i principi fondamentali per cui l'informazione, per essere effettivamente

tale, debba essere reperita, e quindi "letta" e interpretata. *In primis*, dai motori di ricerca!

Una pagina web, per avere qualche chance di visibilità da parte del motore, ha principalmente due possibilità: essere molto popolare oppure farsi notare.

La popolarità di un documento si esprime nel fatto di essere linkato da molte altre pagine, alcune delle quali già inserite nell'elenco delle pagine da visitare. I motori infatti, visitando il Web in maniera ipertestuale – ovvero di link in link – sicuramente prima o poi capiteranno in questa pagina fino ad allora sconosciuta, ma che beneficia di molti altri link su pagine già note al motore. La popolarità – intesa, nel gergo dei motori di ricerca, come quantità di link di cui gode una pagina sul Web – ha una forte influenza nel ranking di una pagina e rappresenta anche una maniera molto efficace per farsi notare dal motore.

L'altra modalità classica per farsi notare dai motori di ricerca è quella di segnalare loro la propria esistenza attraverso i comandi ADD URL/ADD SITE. Basta perlopiù compilare un semplicissimo form in rete in cui si specifica l'URL della pagina e l'e-mail di riferimento, e si ottiene l'iscrizione di una pagina all'elenco delle pagine da visitare da parte dello spider.[10] Il sito potrà essere di conseguenza visitato e sottoposto al processo di indicizzazione solo se sarà effettivamente reperibile online.

Anche questa, che appare come una considerazione di sconcertante banalità, nasconde invece delle problematiche tecniche spesso disattese. Un sito, per essere reperito facilmente dal motore di ricerca, deve risiedere su una macchina in rete ben configurata e in buona salute in termini di funzionamento e banda. L'indirizzo non deve riscontrare difficoltà in termini di indirizzamento;[11] né deve riscontrare problemi di reperibilità sul sito che lo linka. Infatti, se il sito è linkato troppo in profondità in una macchina, il motore che visita la macchina compie un numero troppo elevato di salti di link in link per arrivare al sito stesso, partendo dall'homepage della macchina visitata.[12] Anche problemi di connettività possono ostacolare il reperimento del sito stesso. La macchina in rete che ospita il sito deve garantire una costante presenza online (cosa mai scontata in un sistema complesso come Internet) e con buone possibilità di banda (cosa ancor più difficile da riscontrare) in modo tale che il motore, che periodicamente compie la visita, non incontri intoppi di comunicazione. In relazione ai problemi di connessione, si suggerisce di non produrre pagine eccessivamente pesanti in termini di byte:[13] un'informazione asciutta e leggera ha maggiori possibilità di farsi reperire nelle trame dello spider.

Un altro ostacolo sulla via del reperimento è rappresentato dal caso, molto comune, in cui si cambiano gli indirizzi delle pagine, senza lasciare traccia dei nuovi URL. Quante volte accade che gli utenti di vecchi indirizzi incontrino difficoltà a rintracciare le nuove versioni delle

pagine? Ecco un'altra indicazione che pare ovvia ma viene sistematicamente disattesa dalla maggior parte dei webmaster, anche da coloro che gestiscono i siti istituzionali e importanti!

Ammettiamo, con buona dose di ottimismo, che una certa pagina in rete sia facilmente reperibile dai motori di ricerca; si pone ora il problema di renderla "digeribile" al motore stesso, considerando che il suo metabolismo interpretativo segue regole e algoritmi differenti da quelli usati dai browser di navigazione.

Non è infatti scontato che le pagine web possano essere interpretate correttamente dallo spider del motori di ricerca. Per tentare di comprendere a fondo il problema, immaginiamo che il motore di ricerca legga le pagine di un sito come farebbe un browser di vecchia data, una specie di lynx[14] appositamente configurato per visionare in maniera veloce un'enorme quantità di materiale.

Tanto per cominciare, molti motori di ricerca non leggono tutte le pagine di un sito, ma prendono in esame solo la parte iniziale ("in alto") delle pagine che visitano. Questo vuol dire che viene "tagliato" tutto ciò che si trova al di sotto di questo confine virtuale ("in basso"). Il problema è però nel fatto che quello che il motore intende come linea di demarcazione tra "alto e basso" non necessariamente ha una corrispondenza con qualcosa di visitabile nella pagina. Cosa è "in basso o in alto" nel caso di pagine costruite tramite formattazione a tabelle, molto comuni ultimamente? E nelle pagine a frames? E nelle pagine che contengono in cima javascript, sebbene ai nostri occhi non visibile? Gli occhi del motore di ricerca non sono i nostri[15] e i criteri di percezione non si equivalgono quasi mai. Analogamente, succede nel caso in cui lo spider del motore si imbatta in insiemi di codici o comandi, per esempio i sempre più utilizzati plug-in che permettono la realizzazione di particolari effetti speciali.[16] Inoltre, molti motori di ricerca si trovano in difficoltà quando devono interpretare comandi automatici di reindirizzamento delle pagine come il comune refresh, che spesso viene inteso come un "trucco" per avvantaggiare il ranking di una pagina. Una volta individuata, la pagina che presenta il refresh viene penalizzata dal motore stesso.

Ma anche una pagina semplice, priva di complicazioni tecniche, può risultare difficile da interpretare per gli spider dei motori di ricerca. Prendiamo, per esempio, una pagina web composta unicamente da un collage di immagini anche suggestive. La pagina è dotata di senso sotto il profilo estetico ed artistico per un visitatore umano,[17] ma risulta totalmente priva di significato rispetto a un motore che deve prelevarla e indicizzarla correttamente, dato che non contiene elementi verbali necessari al motore per attribuirgli un senso e riuscire così a indicizzarla correttamente.

In certi casi non solo il motore di ricerca può andare in crisi in presenza di oggetti difficilmente interpretabili (come immagini senza alter-

nativa testuale, plug-in, codice javascript o di altro tipo, frames, comandi di refresh ecc.)[18] ma può anche accadere che il motore reagisca malamente alla presenza di alcune pagine e che le interpreti come "illecite", penalizzandole così come farebbe un rigorosissimo cybercensore.

Ma che cosa è una pagina illecita per un motore di ricerca?

Sono considerate illecite dai motori di ricerca tutte quelle pagine che sono state realizzate appositamente non per offrire un contenuto informativo a chi naviga nel cyberspazio, ma per sottolineare particolari contenuti agli spider. È il caso delle cosiddette "doorway": pagine che ripetono ossessivamente e in maniera considerata efficace rispetto a taluni motori di ricerca, le parole-chiave utilizzate per accrescere la possibilità di un buon posizionamento nelle liste dei motori. Sono pagine molto curiose per la loro pochezza di contenuti e per la ridondanza della scarsa informazione presente e che hanno la proprietà di reindirizzare velocemente, una volta analizzate, il motore sulla pagina vera e propria.[19]

Pagine concepite con questi o altri criteri[20] vengono interpretate da alcuni motori come trucchi escogitati per pompare le pagine in cima alle classifiche, e vengono penalizzate da molti motori. Pagine così concepite non vengono neanche prese in considerazione, oppure inserite in una "lista nera" che le valuta con un grado di attenzione minima. Il motore – purché dotato di automatismi di analisi ben configurati – riesce quindi in alcuni casi a distinguere e penalizzare un atteggiamento considerato illecito.

Nell'ipotesi di progettare un sito con l'obiettivo di renderlo subito popolare dentro, e quindi anche fuori[21] del cyberspazio, e cioè realizzare un sito accessibile ai motori di ricerca, è necessario seguire certi criteri. Ne elenchiamo per chiarezza alcuni tra i più significativi:

1) inventare un nome o un acronimo quanto più originale;[22]
2) acquisire il dominio relativo al nome scelto[23] oppure farsi ospitare da un dominio in cui ci riconosciamo per condivisione di idee o attività professionali;
3) verificare che chi ospita il dominio e il sito garantisca un corretto e costante funzionamento della macchina in rete, in particolare per quanto riguarda le possibilità di banda e connettività;
4) realizzare pagine web originali e creative che, soprattutto, non abbiano caratteristiche ostili ai motori di ricerca;[24]
5) iscrivere il sito manualmente[25] ai motori di ricerca ritenuti interessanti;
6) rendere nota – evitando però di disturbare eccessivamente per non rischiare di diventare notoriamente fastidiosi – la nascita del sito attraverso annunci e iniziative sia in rete sia attraverso altri mezzi di

comunicazione, così da suscitare interesse e forse ottenere anche qualche benefico link da altri siti;
7) utilizzare efficacemente tutti gli strumenti leciti per una corretta indicizzazione del materiale informativo contenuto nel sito: l'attributo ALT di descrizione delle immagini, il TITLE della pagina, i metatag (keyword e description), il testo compreso fra i tag <h1>. In altre parole mettere l'accento soprattutto in cima alla pagina, su cosa si presume appaia allo spider del motore di ricerca controllando che anche nei nomi dei file, delle directory e del dominio appaia la parola-chiave che si vuole sottolineare. E così via.[26]

Come si è potuto intuire non è affatto scontato realizzare un sito, ottimizzando la sua reperibilità e indicizzazione rispetto ai motori di ricerca. Ma perché auspicarsi di essere trovati e ben apparire agli occhi dei motori?

La risposta è semplice: perché desideriamo diffondere al maggior numero di persone le nostre informazioni. Lo scopo finale è sempre la visibilità. E dunque, se finora ci siamo preoccupati di rendere accessibile il nostro ipotetico sito ai motori di ricerca, dobbiamo ora cercare di renderlo accessibile anche alla totalità degli utenti che frequentano il Web.[27]

Come rendere accessibile l'informazione sul Web è una problematica enorme[28] e finalmente emergente. Il pressing caparbiamente effettuato sull'argomento in questi anni da personaggi storici della telematica come Arclele, congiunto all'opera di lobbying da parte di categorie interessate al problema come quelle dei disabili,[29] sta producendo i primi effetti.

Quando questo libro andrà in stampa probabilmente sarà apparsa in Parlamento la prima proposta di legge in materia, mentre iniziative simili già esistono in altri paesi (Francia in testa). I primi segnali[30] di un interessamento politico si sono già avvertiti nell'estate 2000 anche in Italia, in concomitanza con i seminari tenuti all'hackmeeting del CSOA[31] Forte Prenestino, sulla necessità di rendere universalmente accessibili almeno i siti di interesse pubblico. Gli enti pubblici hanno l'obbligo morale e il dovere civile di realizzare siti progettati secondo principi di accessibilità universale, con una presentazione dell'informazione non penalizzante nei confronti di una parte dell'utenza.

Ma che cosa significa rendere universalmente accessibile un sito? Rispetto ai motori di ricerca vuol dire renderlo correttamente reperibile e interpretabile da parte di un qualsiasi spider in cerca di informazioni. Ma significa anche e soprattutto renderlo accessibile a tutti gli utenti possibili – compresi quelli che possono avere difficoltà di varia natura nell'utilizzare strumenti hardware o software.

La preoccupazione di un webmaster è in genere quella di produrre un sito che sia il più colorato, multimediale e intrigante possibile, secondo i canoni estetici di moda su Internet (fortemente sollecitato, in

tal senso, dal committente del lavoro). D'altra parte la fruizione universale di materiale informativo sul Web non segue esattamente questa logica. La creazione di un sito che se offline (per esempio i cd-rom) apparirebbe perfettamente funzionante – ricco di immagini, video, suoni, plug-in flash – è infatti un'operazione legittima che appaga l'occhio e il senso di meraviglia garantendo il consenso del cliente. Tuttavia, lo sviluppo di un sito sulla base di puri canoni estetici non assicura la visibilità del sito stesso sui motori, né tantomeno rappresenta la migliore via per veicolare un'informazione in rete. Innanzitutto, enfatizzare l'aspetto grafico di un sito è una pratica che non ottiene risultati ottimali sui motori, i quali incontrano una serie di difficoltà tecniche per reperire e indicizzare il sito stesso. Inoltre, non rispetta neppure le esigenze della maggior parte dell'utenza non necessariamente dotata di un'ottima connessione hardware e software all'ultimo grido, e si ritrova così a spendere inutilmente tempo e denaro.

Ci sono persone che hanno hardware e software vecchio, di tipo molto diverso oppure diversamente configurato rispetto a quanto auspicato dal webmaster per visualizzare correttamente le pagine. Ci sono ancora persone che hanno il monitor "settato" a 640x480 e non a 800x600. Ci sono persone con scarse risorse macchina e di banda. Ci sono persone che non amano girare con gli active X, i cooky e gli applet Java abilitati perché gelosi della propria privacy. Ci sono persone che non hanno mai scaricato il plug-in flash oppure real-audio perché non ne conoscono l'esistenza o diffidano dei misteriosi messaggi di invito a scaricarli. Ci sono persone che hanno poco tempo per navigare sul Web e quindi non hanno voglia di perdersi in labirinti di effetti speciali. Ma soprattutto esistono persone che possono addirittura essere danneggiate da alcuni movimenti o inaspettate frequenze di movimento degli oggetti ospitati sulla pagina,[32] oppure che non riescono a seguire coscientemente il percorso comunicativo proposto dalla pagina.[33] Esistono persone che non vedono e quindi necessitano di hardware e software di tipo particolare in grado di poter tradurre le pagine web – solo se realizzate secondo alcuni principi di base ispirati all'accessibilità – in testo parlato o Braïlle. Esistono persone che non sentono e che quindi necessitano di traduzioni testuali dell'audio presente su un sito.

Ora, se prendiamo coscienza del fatto che Internet non siamo solo *noi* – il nostro cliente e qualche persona con la quale occasionalmente chattiamo – e comprendiamo che il Web è un potenziale infinito di risorse tecniche e umane di cui è impossibile conoscere le caratteristiche, allora forse non ci sarà difficile concordare che esiste in rete un universo umano e tecnologico che attualmente sfugge alle logiche di marketing, ma che non può e non deve sfuggire alla nostra coscienza e sensibilità.

La libertà di una singola persona, insomma, è la libertà di tutti.

Questa parola d'ordine è una delle tante che cercano di descrivere la

complessità di vivere la solidarietà in senso universale, anche se spesso viene interpretata solo in termini restrittivi e talvolta instilla niente altro che un atteggiamento caritatevole. Eppure dovrebbe comportare qualcosa di più profondo perché ciò che è universalmente accessibile, ciò che è utile per le persone che hanno meno risorse, è in realtà utile per tutti![34]

Realizzare una pagina web accessibile[35] è un vantaggio per tutti, soprattutto in determinate situazioni avare di connettività. Costruire una pagina che offra in maniera spartana, ma velocemente scaricabile, un contenuto informativo chiaro e puntuale, rappresenta una risorsa di grande importanza. Se poi tutto ciò si dovesse scontrare comunque con necessità aziendali che impongano di realizzare pagine attraenti ma molto pesanti, si dovrebbe proporre di progettare una doppia versione del sito: accanto alla versione high-tech, ce ne sarà una universalmente accessibile.

Le strategie tecniche[36] da adottare per realizzare una pagina ad accessibilità universale diventano quindi termini di una rivendicazione politica.[37] Accanto ai criteri indispensabili per massimizzare la visibilità per i motori di ricerca, indichiamo quindi altri utili suggerimenti per costruire una pagina accessibile:

- evitare di usare codice proprietario[38] non conforme agli standard del W3C;[39]
- evitare di ottimizzare le pagine solo per una data versione di browser o risoluzione video;[40]
- cercare di suddividere la pagina in sezioni chiaramente riconoscibili in modo da ottimizzare la ricerca delle informazioni sul sito; cercare di distinguere fra struttura e presentazione;[41]
- valutare attentamente i tempi di caricamento delle pagine,[42] considerando che spesso gli utenti che visitano un sito sono alla ricerca di qualcosa di preciso, non conoscono la struttura del sito, non hanno molto tempo da dedicare alla visione della pagina, non amano i lunghi tempi di attesa nel download;
- in presenza di link impiegare un testo che abbia senso anche al di fuori del contesto limitato della pagina e in una lettura che avviene in successione;[43]
- non scrivere documenti aventi linee che superino le 80 battute per riga;
- per le tabelle[44] utilizzare valori relativi in percentuale (%) e non assoluti; utilizzare ascii-standard;[45]
- garantire un buon contrasto tra testo e sfondo, assicurarsi che testo e grafica siano percepibili e comprensibili anche se visualizzati in assenza di fogli di stile (CSS), comunque non utilizzare colori per trasmettere contenuti informativi;

– non ricorrere a oggetti, comandi, applicazioni o tag che provocano movimenti inaspettati delle pagine;
– identificare i linguaggi naturali utilizzati nel documento con gli appositi marcatori,[46] specificare lo scioglimento di ogni abbreviazione o acronimo nel documento laddove compare per la prima volta, e tentare di tradurre la pagina in più lingue;
– se si vuole o si deve aggiungere un oggetto non testuale nella pagina[47] prevedere un metodo alternativo di accesso[48] al contenuto di quell'oggetto e una sua descrizione testuale;[49]
– nel caso sia presente nella pagina un audioclip oppure un video, provvedere a fornire un commento chiaro dell'oggetto con l'attributo ALT e inserire un testo scritto del parlato a uso delle persone prive di udito;
– nel caso di video, inserire un file audio in cui un narratore illustri cosa avviene nel filmato in maniera sincronizzata, ad uso di persone ipovedenti;[50]
– controllare il lavoro svolto con strumenti di valutazione.[51]

NOTE

1 Borges, nel 1941, nella *Biblioteca di Babele* scriveva: "L'universo (che altri chiamano la Biblioteca) si compone di un numero indefinito, forse infinito, di gallerie esagonali... Quando si proclamò che la Biblioteca comprendeva tutti i libri, la prima impressione fu di straordinaria felicità. Tutti gli uomini si sentivano padroni di un tesoro intatto e segreto. Non v'era problema personale o mondiale la cui eloquente soluzione non esistesse in un qualche esagono. L'universo era giustificato, l'universo attingeva bruscamente le dimensioni illimitate della speranza... Alla speranza smodata, com'è naturale, successe una eccessiva depressione. La certezza che un qualche scaffale d'un qualche esagono celava libri preziosi e che questi libri preziosi erano inaccessibili, parve quasi intollerabile... Sappiamo anche di un'altra superstizione di quel tempo: quella dell'Uomo del Libro. In un certo scaffale d'un certo esagono (ragionarono gli uomini) deve esistere un libro che sia la chiave e il compendio perfetto di tutti gli altri: un bibliotecario l'ha letto, ed è simile a un dio... Come localizzare l'esagono segreto che l'ospitava? Qualcuno propose un metodo regressivo: per localizzare il libro A, consultare previamente il libro B; per localizzare il libro B, consultare previamente il libro C; e così all'infinito... In avventure come queste ho prodigato e consumato i miei anni." Secondo Ernesto Hofmann (*Un'architettura per l'informazione digitale*, "IF", 1999, n. 1) in *Finnegans Wake*, "Joyce sembra quasi anticipare l'infostruttura del cyberspazio, nel quale il flusso delle informazioni avviene attraverso continue scariche di entità che, entrando nel sistema attraverso una 'portal vein' (che ricorda i 'portal' di Internet!), vengono distribuite, dissociate e ricombinate. Joyce che, in *Finnegans Wake*, più volte parla di televisione (e siamo solo negli anni Trenta), era profondamente affascinato dalla sempre più complessa infostruttura elettronica che si stava evolvendo. Il formidabile "chaosmos of Alle" (118, 21) che si andava configurando gli faceva intravedere un'unica memoria per una quantità pressoché infinita di informazioni. Che ne sarebbe stato allora del libro?

Forse, proprio in contrapposizione al prossimo avvento delle realtà virtuali del cyberspazio, andò a cercare i suoi mondi virtuali all'interno del linguaggio umano servendosi della lingua ambigua e spesso allucinata dei sogni."

2 Ovviamente è da intendersi per "scelta" la configurazione adottata dal motore di ricerca per un suo corretto funzionamento, relativamente alle attese del programmatore, o meglio dell'azienda che commissiona il lavoro. Tali configurazioni sono cambiate frequentemente e spesso non sono neanche esplicitate in maniera trasparente, tanto che l'unica maniera per tentare di individuarle è quella di operare una scrupolosa opera di reverse-engineering su questi software di ricerca.

3 Dalla *Dichiarazione Finale dell'ICATA 89* adottata il 4 agosto 1989: "Ogni informazione è al contempo deformazione. Il diritto all'informazione è al contempo inesorabilmente legato al diritto alla deformazione, che appartiene a tutto il mondo. Più si produce informazione, e più si crea un caos di informazione, sfociante sempre più in rumore. La distruzione dell'informazione come del resto la sua produzione, è il diritto inalienabile di ognuno".

4 Non tanto per le nostre capacità tecniche, ma piuttosto perché condividiamo la valenza sociale dell'abbattere le barriere presenti nel mondo della comunicazione e dell'informazione.

5 Per approfondimenti in materia si consiglia di ricercare documenti aggiornati e facilmente reperibili in rete dei tanti studiosi che seguono e alimentano il dibattito scientifico in materia come Gerald Salton, Linda Schamber, Nicholas Belkin, Bruce Croft, Peter Foltz, R. Wilienski e Giuseppe Attardi.

6 Vedremo più avanti alcuni esempi di applicazione di questo concetto.

7 All'elenco delle stopword appartengono anche quelle parole la cui frequenza sia al di sotto o al di sopra di una certa soglia prestabilita; in altre parole, questo elenco comprende i termini troppo rari o al contrario troppo frequenti.

8 I tag contrassegnati con la lettera h1 sono quelli che all'origine di Internet e secondo il linguaggio HTML che genera le pagine web – la maggior parte dell'informazione presa in considerazione dai motori di ricerca – caratterizzavano il testo come molto importante a livello strutturale.

9 Anche il WML, il linguaggio per creare pagine ipertestuali correttamente interpretabili dai cosiddetti "telefonini", ha molto in comune con le priorità di un documento accessibile: questo perché – e sarà vero anche per i nascenti apparecchi UMTS – uno strumento palmare visualizza bene delle pagine che sono realizzate per risorse – quanto meno di visualizzazione – minime.

10 Questo non funziona necessariamente nei casi in cui esista una redazione che stabilisce il grado di interesse – secondo i suoi parametri – di un sito segnalato alla directory.

11 Ovvero per esempio in termini di configurazione dei relativi DNS – Domain Name Server.

12 Stesso discorso ovviamente vale per l'interno del sito. Qualora il sito sia troppo profondo o con troppi link concatenati fra loro, o, viceversa, con un numero troppo basso in maniera tale da avere scarsa connettività, possono verificarsi problemi di reperimento di parte del sito.

13 Una vecchia regola della rete dice di non produrre documenti che siano pagine web ma anche messaggi e-mail – non più grandi di 32 Kb.

14 Storico ma efficientissimo browser testuale reperibile su http://lynx.browser.org/

15 Per credere provare il Search Engine Simulator reperibile su http://www.delorie.com/web/ses.cgi. L'incommensurabilità tecnica fra "i sistemi di visione" utilizzati dai browser per leggere il Web crea una discrepanza netta: il Web visto da lynx non equivale al Web visto da Netscape o da altri comuni browser.

16 Con il risultato di scaricare pezzi di testo inutili o fuorvianti per l'indicizzazione o peggio ancora che possono confondere lo spider inducendolo a non effettuare lo scaricamento completo del sito.

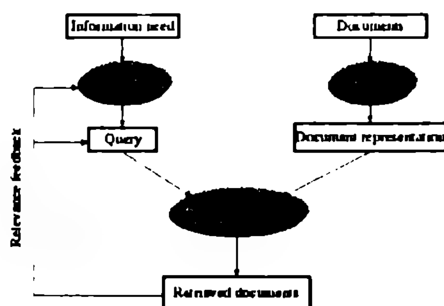
17 Normodotato.
18 Ovviamente, non è sempre vero per tutti i motori di ricerca, alcuni dei quali sono in grado di interpretare correttamente una pagina-frame ma reagire negativamente alla presenza del comando REFRESH.
19 Per esempio con il già citato tag REFRESH, oppure con un codice javascript funzionale allo scopo.
20 Citiamo per esempio il "trucco" di mettere del testo in cima alla pagina dello stesso colore dello sfondo in maniera tale da servire a un'eventuale indicizzazione del motore, ma risulti invisibile all'occhio umano.
21 Non dimentichiamo mai che quello che facciamo in rete ha sempre una ricaduta – seppur indiretta – sulle coscienze e sulle realtà fisiche che ci circondano quotidianamente.
22 Se si sceglie un nome o un acronimo inflazionato e usato in situazioni già molto conosciute, sarà ancor più difficile farsi notare! La scelta è tanto più efficace quanto più creativa si mostra.
23 Esistono innumerevoli servizi a riguardo, alcuni anche molto economici, come lo è attualmente www.DomainMonster.com.
24 Si consiglia caldamente a questo scopo di consultare le relative pagine di help presenti sulla maggior parte dei motori di ricerca.
25 Si consiglia di iscrivere manualmente il sito ai motori di ricerca sia perché in tal modo diventa possibile specificare eventuali informazioni richieste dai singoli motori, sia perché alcuni software automatici di iscrizione possono non effettuare il lavoro così come dichiarato o lo fanno in maniera tale da indispettire il motore di ricerca stesso, il quale può interpretare alcune opere di intensificazione di iscrizione come spam e quindi penalizzare il sito iscrivente.
26 Per approfondimenti vedi http://strano.net/chaos considerando, comunque, che anche per questo particolare ambito telematico le regole sono in continua evoluzione. Una paziente, periodica e aggiornata opera di reverse-engineering sui motori di ricerca appare comunque l'unico strumento utile da consigliare in materia. Se volete tentare di apparire in alto rispetto al tema del birdwatching cercate di diventare popolari in materia con il sito, ma verificate anche le caratteristiche dei siti che appaiono nel top-ranking dei motori di ricerca digitando appunto birdwatching.
27 Parlando di accessibilità non possiamo fare a meno di ricordare che la maggior parte delle aree geografiche – Africa *in primis* – è tagliata fuori non solo dal circuito telematico, ma anche da quello più semplicemente telefonico (cfr. il tema del *digital divide*).
28 Per approfondimenti tecnici sull'argomento si consiglia di visitare periodicamente il portale sull'accessibilità xs2web (Access To Web) ~ http://www.ecn.org/xs2web.
29 In particolare gli ipovedenti.
30 L'AIPA – Autorità per l'informatica nella pubblica amministrazione – ha avviato un'indagine esplorativa su come poter rendere maggiormente accessibili gli strumenti hardware e software dei siti della Pubblica amministrazione, mentre si sono registrate numerose interpellanze politiche in materia provenienti da partiti politici di vario segno.
31 Centro sociale autogestito.
32 È il caso delle persone che soffrono di epilessia.
33 È il caso di persone con disabilità cognitive, ma anche di molte persone cosiddette "normodotate" che restano facilmente frastornate da così tanta informazione accumulata in maniera indecifrabile.
34 La condivisione e socializzazione delle risorse è una pratica utile a tutti.
35 Non bisogna poi dimenticare le barriere linguistiche, per cui sarebbe necessario fornire il sito tradotto in più lingue.
36 Di nuovo, anche sotto il profilo tecnico, suggeriamo www.ecn.org/xs2web come riferimento manualistico in materia.

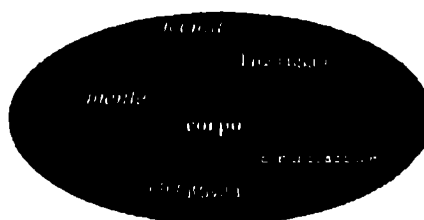
37 Cercheremo di essere brevi anche se l'argomento è enorme. È infatti sicuramente possibile realizzare pagine che siano universalmente accessibili per qualsiasi tipo di utenza, ma è pur vero che è possibile anche un tipo di "politica graduale" in materia. Per esempio è consigliabile non realizzare pagine con frame, ma è anche vero che dovendo necessariamente costruire una struttura a frame, è possibile realizzarla secondo criteri più o meno accessibili. Possiamo quindi parlare di accessibilità parziale, anche nel caso in cui venisse definitivamente approvato l'HTML4, che prevede una serie di migliorie sul tema dell'accessibilità, ma che non potrà tener purtroppo conto di coloro che continueranno a navigare con browser che non riconoscono le regole dell'HTML4.
38 Per esempio, Word ecc.
39 Il World Wide Web Consortium ha elaborato a tal merito delle regole specifiche reperibili in rete all'indirizzo http://www.w3.org/WAI.
40 La definizione che normalmente viene usata dagli ipovedenti.
41 Organizzare cioè l'informazione e la struttura del sito in maniera tale che l'utente si possa muovere al suo interno agevolmente.
42 Non utilizzare più immagini di quante non ne servano, e cercare di riutilizzare più volte le stesse immagini fra/nelle diverse pagine del sito, limitando l'uso di immagini ad alta risoluzione che appesantiscono la pagina aumentando i tempi di download.
43 Per esempio, non usare "clicca qui" e tenere conto che per alcune categorie di utenti la lettura del testo e quella dei link risulta separata: è quindi importante che la lettura in successione dei link abbia un senso. Per la stessa ragione è importante evitare di mettere link adiacenti e, nel caso questo sia necessario, separarli e delimitarli utilizzando spazi e caratteri di testo. Una buona simulazione di questa situazione si ottiene ripetendo il tasto TAB in una pagina visitata dal browser.
44 Cerchiamo di non utilizzare tabelle per formattare il testo; piuttosto identifichiamo le intestazioni di righe e colonne attraverso i marcatori che associano le caselle di dati alle celle di intestazioni; inseriamo un link a una descrizione letterale dei dati inseriti in tabella. I dati in forma di tabella possono essere rappresentati anche tramite testo preformattato utilizzando l'accortezza di impiegare spazi e non tabulazioni, poiché quest'ultime variano da client a client.
45 Vocali non accentate ecc.
46 In questa maniera, per esempio, traduttori vocali o braille sanno di dover impiegare i relativi dizionari per interpretare la pagina al meglio delle loro possibilità.
47 Una mappa cliccabile, un'immagine, uno script, un applet, un'ascii-art ecc.
48 Per esempio attraverso i tag specifici NOFRAMES e NOSCRIPT.
49 Attraverso i tag ALT e TITLE e un (D) link, ovvero link a una pagina di solo testo che descrive l'oggetto individuato da una "D".
50 In entrambi i casi fornire sempre l'equivalente testuale.
51 Uno per tutti, il più severo: http://www.cast.org/bobby.

Durante la fase di ricerca nei motori, il gioco linguistico e l'approccio creativo sono spesso componenti necessari per ottenere informazioni più precise e allo stesso tempo fattori determinanti per l'indicizzazione dei dati e la costruzione di database.

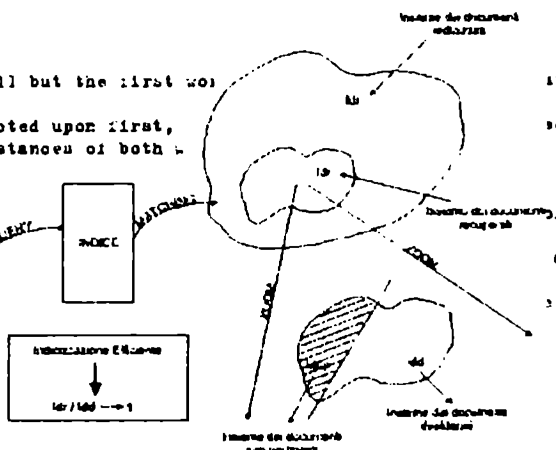
xs4all
ACCESSO PER TUTTI

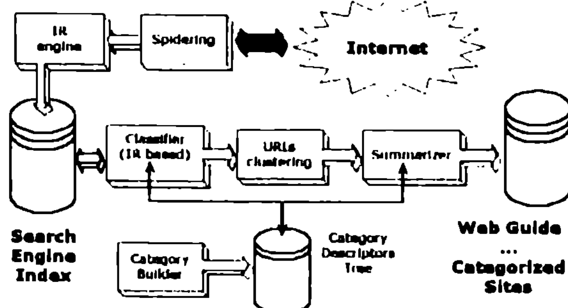
The retrieval process

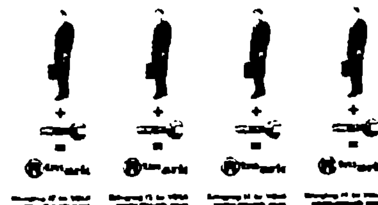




Expression Comment cars AND bmw AND mercedes
cars & bmw & mercedes
Finds documents that contain each of these words. All but the first wo
Sort by: bmw mercedes
As in Algebra, expressions within parentheses are acted upon first,
Results are sorted so that pages containing more instances of both
Mark McGwire
Sort by: baseball "home run" cardinals
Finds pages with the phrase Mark McGwire and list
Sort by: engineering MS nuclear
Finds pages with the word resume on the page or i
Sort by: "nuclear power plant""
Limits the search to the domain gov and then look
Notice that this search combines a command (domai
golf AND Monterey AND url:golf course Limits the :









6

LINGUISTICA

6.1 ORIZZONTI LINGUISTICI

> "Saussure insiste con forza sul carattere potenziale, sulla 'productivité' e, come egli dice, sulla 'créativité' della lingua: il fatto che una determinata combinazione sintagmatica esista ha una importanza assai minore del fatto che essa *possa* esistere".
> Tullio de Mauro, *Introduzione al Corso di linguistica generale*, di F. de Saussure

Le implicazioni della linguistica in relazione alle metodologie di indicizzazione del materiale in rete sono ancora poco esplorate, così come poco battuti sono i territori che propongono[1] l'analisi e la trattazione del materiale testuale attraverso i metodi statistici, sintattici e logici. Perché allora addentrarsi in questo campo minato? In realtà questo è un argomento centrale che mette in luce, secondo l'ottica del gioco linguistico, la funzione fondamentale di qualsiasi risorsa di ricerca e chiarisce le pratiche attuate dagli utenti per indicizzare le loro pagine web.

I motori di ricerca hanno la funzione e il compito di soddisfare il più possibile le aspettative dell'utente. Le risposte relative a una certa richiesta devono essere sempre più efficaci e precise; il motore di ricerca deve fornire un'elevata velocità di reperimento dei dati, senza perdere

in precisione.[2] Dal punto di vista dell'utente, non in qualità di ricercatore ma come "imputter" di dati che vengono indicizzati da parte delle risorse di ricerca, bisogna invece prendere in considerazione le tecniche e strategie sintattico-linguistiche che consentono ottimi riscontri orientati al ranking.

Per visualizzare la questione in senso complessivo, occorre di nuovo evidenziare la netta distinzione esistente tra i due primari strumenti di ricerca attuali: il motore di ricerca e la web directory.

Il motore di ricerca si basa su principi automatici: utilizza gli spider che cercano in Internet le pagine per la costruzione del database consultabile dall'utente attraverso l'uso delle parole chiave. La web directory, sorta di catalogo realizzato con il lavoro redazionale umano, richiede invece una navigazione manuale, consente un'estrazione delle informazioni ritenute più significative, e presuppone un'organizzazione tassonomica dei contenuti. Molti dispositivi di ricerca usano entrambi i metodi di indicizzazione creando un mixer tra soluzione software e soluzione umana, ma l'incremento esponenziale di Internet impone sempre più una riduzione della seconda componente a favore dell'applicazione di agenti intelligenti.

A livello linguistico, i due differenti modelli di ricerca richiedono che il motore di ricerca garantisca una completa sistematicità e permetta un'indicizzazione a vasto raggio dei documenti esistenti in rete. Il risvolto negativo di ciò è l'aumento del grado di incertezza e ambiguità del valore semantico delle unità lessicali indicizzate.[3] La web directory risolve, in parte, il problema dell'attribuzione del significato delle parole garantendo più ordine e coerenza, anche se ciò comporta una minore quantità di materiale indicizzato e provoca ritardi nell'aggiornamento dei risultati consultabili. Lo slogan per le web directory è "indicizzare il meglio", ma la lentezza degli operatori umani a confronto degli agenti software, sia in fase di primo recupero dell'informazione sia durante le successive visite ai web site (sottoposti a continui aggiornamenti), penalizza la web directory sotto il profilo quantitativo.

Poiché il problema della categorizzazione degli elementi linguistici riguarda in maniera minore le web directory, l'oggetto di questa analisi si focalizzerà unicamente sull'attività svolta dai software. Le posizioni interpretative stabilite dagli operatori umani all'interno di una redazione, se da un lato possono risultare più o meno discutibili, dall'altro certamente diminuiscono le possibilità di errore nell'attribuzione di significato degli elementi linguistici, errore che invece risulta più evidente nelle azioni degli agenti intelligenti.

6.2 IL "SIGNIFICATO" NEI MOTORI DI RICERCA

Nelle ricerche in rete, una delle principali funzioni dei motori è quella di attribuire un significato non ambiguo alle unità linguistiche, siano

esse minime (monemi e morfemi) o complesse, quali le sottounità (suffissi, prefissi e radicali) costituite da catene più elementari.[4] Tenendo conto dell'abnorme incremento quantitativo dei dati testuali in rete, la loro dinamicità e la diversificazione della natura dei testi: testi scientifici, tecnici, poetici, legislativi, normativi, corrispondenza ecc.

Una stessa idea può essere espressa da combinazioni multiple di differenti parole e costruzioni sintattiche. Si pensi, per esempio, all'uso dei sinonimi che, se da una parte arricchiscono un testo permettendo di evitare la ripetizione degli stessi termini, dall'altra creano delle complicazioni tanto nelle fasi di recupero delle informazioni, costringendo gli spider a un maggior lavoro, quanto nella ripetizione nel testo delle keyword da promuovere.[5]

Si apre così la discussione sulla ricerca in rete basata sulle parole chiave: un modello che fa emergere la difficoltà di formulazione della richiesta e l'inerente ambiguità della parola, in senso lato, nell'ambito del linguaggio naturale. Il problema viene chiaramente illustrato dallo scenario della ricerca di informazioni sul Web, dove gli utenti si sentono fortemente impacciati di fronte al form della query di qualsiasi interfaccia di un motore di ricerca: è a volte problematico formulare richieste complesse che usualmente si riducono alla composizione di una sola parola o coppia di parole. Ma il percorso è ancora lungo e irto di difficoltà, perché per ogni richiesta il numero di risposte che appaiono nelle hit è altissimo, generando confusione e straniamento nell'utente stesso. Le ragioni di questo fatto sono da attribuire alla relazione tra identità e differenze delle parole nel linguaggio naturale, e alle varianti interpretative delle richieste da parte dei motori di ricerca. Cioè, formulando le stesse espressioni di interrogazione, riceviamo responsi concettuali diversi che dispongono su diverse prospettive linguistiche l'informazione risultante.

Possono essere escogitati da parte dell'utente, dei rimedi a questo rilevante limite dell'*information retrieval*[6], il quale deve elaborare una strategia mirata, sforzandosi di usare query complesse e di giocare soprattutto con le espressioni linguistiche, in particolare usando la ricerca avanzata offerta dai maggiori motori di ricerca. Può operare una selezione scegliendo in maniera chirurgica l'informazione reale desiderata, e restringendo il campo della domanda. Ma anche il motore di ricerca, attraverso un miglioramento del feedback, ottenuto allargando la concettualizzazione dei risultati, può a sua volta produrre un'espansione lessicale-semantica del significato delle parole. Alcuni ricercatori a questo proposito stanno lavorando nella direzione proprio dello sviluppo di software, algoritmi, agenti intelligenti capaci di avvicinarsi sempre più alle esigenze dell'utente.[7] La macchina che "capisca" quali sono le intenzioni linguistiche dell'uomo: ecco l'obiettivo da raggiungere! Interpretare il linguaggio naturale, nel mondo linguistico-informatico,

dovrebbe significare "parlare al motore di ricerca" semplicemente come farebbe con una persona qualsiasi che intende la nostra stessa lingua. La ricerca di punta si sta dirigendo verso la progettazione di un motore di ricerca che vada al di là della logica booleana e che ragioni seguendo la logica fuzzy,[8] uno strumento di recupero delle informazioni al quale non debba più essere rivolta un'interrogazione come per esempio – (dipinti near tempera) and "galleria d'arte", ma "dove posso comprare un dipinto realizzato a tempera".[9]

6.3 NEL CUORE DELL'INDICIZZAZIONE

L'indicizzazione dei documenti è l'operazione che permette di costruire, in modo ordinato e semplificato, l'indice del patrimonio di dati che viene messo a disposizione degli utenti che desiderano effettuare le ricerche in rete.[10] Per il recupero delle informazioni i motori di ricerca si avvalgono di metodi e processi di indicizzazione progettati sin dagli anni Sessanta,[11] e che sono oggetto di continui studi e sperimentazioni al fine di estenderne le potenzialità.

I vari metodi di indicizzazione perseguono l'obiettivo primario di costruire una specie di dizionario, interno al motore di ricerca, che risponda alle diverse esigenze degli utenti. Come abbiamo visto nel precedente capitolo (par. 5.1) la fase di indicizzazione si basa sul lavoro di analisi linguistica delle svariate tipologie di informazione. Un momento essenziale del processo di indicizzazione delle parole è lo *stemming*. Determinare le *stems*, cioè identificare l'insieme di parole che posseggono la stessa radice (*stem*=radice), consiste nel radicalizzare tutti quei termini che pur possedendo differenti significati, possono essere ricercati velocemente grazie al fatto di comprendere la medesima radice; per esempio: arte-artista-artistico-artefatto-ecc.[12] Questo procedimento di riduzione a elementi-origine è molto complicato e necessita accortezza, non tanto nel caso dell'esempio citato (art), nel quale la radice mantiene quasi interamente il suo valore concettuale anche quando è parte di altri termini, quanto nei casi in cui una stessa radice identifica termini con significati completamente diversi. Si pensi, per esempio, alla radice port che accomuna porto-portinaia-portata-portale.

Un metodo di indicizzazione, che cerchi di affrontare questi tipi di inconvenienti, anche se difficilmente implementabile, è il Concept Index che in ogni radice individua una serie di termini semanticamente contigui e li dispone all'interno di una mappa semantica capace di rendere conto anche del grado di contiguità tra i termini associati.

Anche con questo metodo sorgono dei problemi quando ci riferiamo a materiale testuale di carattere generale, difficile da contestualizzare. Un grosso aiuto ai problemi che sorgono nelle fasi di indicizzazione viene dalla realizzazione di tesauri che raccolgano sinonimi, oppure di diziona-

ri con frasi esatte, anche se sicuramente in futuro ci si orienterà sempre più verso la realizzazione di applicazioni sviluppate nell'ambito dell'intelligenza artificiale.

I diversi approcci metodologici attuali, le riflessioni teoriche e le ricerche tecniche che riguardano l'indicizzazione fanno comunque tutti riferimento a un concetto essenziale: la pertinenza. Purtroppo, al fine di stabilire dei criteri generali di pertinenza, i motori spesso tendono a ridurre e limitare la portata semantica dei termini, ingabbiando i termini stessi all'interno di spazi semantici impoveriti. Per esempio, i motori di ricerca rendono un "oggetto" pertinente operando secondo il principio di esclusione (la parola ripetuta o insignificante è eliminata) oppure in virtù di regole correlative e di similitudine (i segni che producono gli stessi concetti vengono raggruppati). Questo funziona, naturalmente. Ma limita anche la possibilità di un approccio più creativo e ludico alla rete e alle operazioni che in rete e con la rete si possono compiere.

6.4 IL SAPERE COME GIOCO LINGUISTICO

Sappiamo che un sito acquista visibilità quando viene correttamente gestita una serie di operazioni (*tips and tricks*) che tendono a massimizzare la pertinenza tra i termini cercati e i risultati trovati, mirando a spendere quante meno energie e risorse possibili. L'aumento dell'output, nel senso di risultati ottenuti, deve corrispondere a una diminuzione dell'input, nel senso di energia spesa. La costruzione del top ranking è dunque un'attività imparentata al minimalismo e alla sottrazione:[13] un metodo che raggiunge ottimi risultati restringendo il campo dell'azione e della creatività. Dal momento che ogni operazione di ricerca è unicamente vincolata all'ottimizzazione della visibilità rispetto agli spider, finisce per risolversi in un'attività di sottrazione linguistica, fortemente tesa a conservare l'insieme di regole utili alla gestione della ricerca piuttosto che a promuovere un'utopistica ricerca sperimentale. Al contrario, un approccio creativo (o "rivoluzionario") consiste in una prolifica–provocatoria–ludica ricerca innovativa che vuole stravolgere le regole del gioco e si contrappone con forza all'atteggiamento conservatore che propone una irriflessiva accettazione delle mosse ritenute proficue. Questa pratica è quindi imperniata su un lavoro più profondo e sotterraneo:[14] il tentativo di creare *ex novo*, di contaminare e di produrre/deformare il senso attraverso la ricerca di neologismi. In altre parole, cortocircuitando l'ortodossia grammaticale in modo tale da piegarla e dirottarla verso il territorio sperimentale dell'immaginazione linguistica.[15] Si tratta di una pragmatica[16] che libera le energie occultate dal "conservatorismo linguistico", che a sua volta promuove solamente il "gioco del top ranking" e sviluppa flussi magmatici dagli imprevedibili e disinibiti risvolti linguistici, consentendo sia di introdurre nuove mosse linguistiche, sia di cambiare

le regole del gioco.[17] Solo il *deturnamento* linguistico, nel senso situazionista di Debord,[18] lo sperimentalismo semantico più sfrenato possono infatti generare un'autentica evoluzione nei giochi linguistici – in particolare nel campo dell'*information retrieval.*

6.5 LA PRAGMATICA RIFORMISTA. L'ESEMPIO DEL TOP RANKING

"La pragmatica riformista" provoca un impoverimento dei giochi linguistici e ingenera una logica di costrizione e involuzione linguistica. In questo caso il miglioramento delle prestazioni, misurabile in termini di visibilità del sito, significa anche una limitazione delle caratteristiche formali e delle soluzioni realizzative. Tutto ciò che non è funzionale finisce nel cono ombra della rimozione forzata.

Ma ci sono altre innegabili e preoccupanti conseguenze. Questa pratica produce una subdola equazione fra visibilità in rete e il binomio efficienza-ricchezza. Senza risorse economiche non è attualmente possibile finanziare uno studio specifico sulle più aggiornate tecniche del ranking, studi spesso affidati ad agenzie di settore che praticano costi elevati. In altre parole: senza ricchezza niente tecnica; ma è anche vero il contrario: niente ricchezza senza tecnica.[19] Infatti, chi pur operando in rete non risulta visibile agli occhi dei motori di ricerca, di fatto non esiste, e a sua volta ha scarse possibilità di esistere nell'impero dell'e-business. I motori di ricerca quindi si trasformano: da strumenti tecnologici di ricerca dell'informazione, diventano un mezzo per la circolazione del capitale; nell'offrire visibilità a chi è in grado di escogitare soluzioni tecniche sempre più forzate e corazzate, rendono la ricerca dell'informazione un prodotto in vendita nelle vetrine virtuali dell'e-commerce. Il motore di ricerca, finalizzato alla sola performatività, si trasforma in una forza "produttiva" la cui principale sfera d'azione è sempre più rivolta al mondo della pubblicità e del business. Basti pensare alle campagne di marketing tramite l'utilizzo di banner nelle pagine dei motori, o ai programmi di partnership[20] e ai lavori tecnici delle agenzie di top ranking.

L'amministrazione della ricerca dell'informazione cade così sotto il controllo di una logica unicamente orientata all'efficienza. Non si ricerca informazione né si dà accesso all'informazione per soddisfare "sete di conoscenza"[21] ma per accrescere la potenza.

Il gioco del ricercare in funzione della piena efficienza nulla ha a che spartire con l'invenzione linguistica o lo sperimentalismo concettuale, ma è piuttosto un gioco basato sulla forza. Non è più la forza concentrata a eliminare direttamente la concorrenza con atti di censura bruta come all'epoca dell'Indice dei libri proibiti.[22] Ma è invece una forza di natura economica e tecnologica che cerca di imporsi sugli avversari operando con strategie tutte orientate al successo del criterio performativo-tecnologico. So-

no mosse studiate a tavolino fin nei minimi dettagli, che dettano legge nelle normative vigenti in nome della presunta efficacia delle procedure.[23]

È possibile rispondere a un simile attacco tecnologico? Esiste una strategia alternativa che, giocando con le possibilità linguistiche e la creatività artistica, sia in grado di fronteggiare l'assalto di una tecnocrazia sempre più solida? La via "rivoluzionaria" della pragmatica creativa è una delle risposte più stimolanti.

Pur senza essere in grado di spiazzare in modo radicale lo strapotere di chi tenta di arrogarsi la fetta di maggior visibilità del Web,[24] la pragmatica creativa riesce a stravolgere i (con)sensi costituiti sulla base di regole canoniche e canonizzate, limitando il controllo tecnocratico dell'informazione e tentando di garantire un buon livello di pluralismo dell'informazione stessa.

6.6 LA VIA RIVOLUZIONARIA O PRATICA CREATIVA

La pragmatica creativa si fa portavoce di un pensiero emancipativo che caratterizza tanto le piccole realtà storiche di Internet quanto le realtà emergenti che, pur prive di risorse economico-finanziarie, sono dotate di una forte carica immaginativa e di carisma cibernetico. Per esempio il sito olandese di xs4all ha fatto scuola. Si tratta di un sito presente su Web dal 1995, il cui stesso nome richiamava l'utopia dei primi cibernauti che intendevano fornire accesso gratuito alla rete per tutti[25] e che fin dall'inizio ha cercato di dare grande risalto all'immaginazione, lavorando non soltanto sui contenuti, ma anche sul linguaggio, come il nome stesso del sito lascia intendere. Il nome, basato sull'assonanza linguistica, sulla sonorità dell'accostamento fra lettere alfabetiche e numeri che scaturisce dalla pronuncia inglese dell'espressione xs4all – *Access For All* – rompe innanzitutto la logica performativa appena discussa e resiste alle tentazioni di legittimazione tecnocratica di chi invece si affida a trucchi di bassa lega... Gli hacktivisti di xs4all,[26] come di Isole nella rete,[27] giocano invece con le strategie subliminali[28] e nei loro ricchi e ludici giochi linguistici privilegiano la fonetica[29] alla semantica creando un deturnamento concettuale e uno spiazzamento semantico-intellettuale. Mutuando un armamentario di tecniche retoriche dalle poetiche dell'underground beat e non solo (dal *cut-up* di Burroughs ai paradossi linguistici di Lewis Carrol,[30] dalle tecniche più sofisticate di scrittura automatica elaborate da Joyce alle punkzine di fine anni Settanta),[31] la strategia dell'hactivism ha in effetti ricreato in rete un accentuato sperimentalismo presente per esempio nel cyberjacking di *etoy*.[32] È un impiego della lingua per il gioco linguistico dalle enormi potenzialità. La sperimentazione permanente citazionista – plagiarista – o cyberpunk nasce per arricchire disordine vitale all'informazione online, che tutto deve essere tranne

che cristallizzata in stereotipi legati alla performatività e alle esigenze strettamente commerciali.

Se il principio di efficienza produce l'effetto di subordinare l'informazione (e a sua volta la ricerca dell'informazione su Web) a un principio di potenza, la pragmatica creativa rivoluziona questo stato di cose generando T.A.Z., zone di autonomia temporanea di forte impatto immaginativo.[33]

La pragmatica creativa non si attua però soltanto nel caso dell'inserimento di informazione in rete, e dunque non si esaurisce nella creazione originale di acronimi, siti, contenuti dalle connotazioni "free". Il percorso "rivoluzionario" di questa pratica si rivela anche nell'*information retrieval*, poiché nelle sue linee guida di libertà espressiva indica la via per una ricerca libertaria e autogestita dell'informazione. In particolare suggerisce quello che Lyotard già definì "l'uso raffinato del gioco linguistico interrogativo": l'interrogazione creativa. Per aggirare i rischi di tecnocontrollo bisogna imparare a indirizzare le domande, formulandole in modo tale da non finire necessariamente nelle subdole maglie del ranking. "La domanda che bisogna porsi nell'ambito di questa propedeutica non è più: è vero? ma: a che cosa serve? Nel contesto della mercificazione del sapere, tale domanda significa nella maggior parte dei casi: si può vendere? E, nel contesto dell'incremento di potenza: è efficace?". Per evitare le "politiche di seduzione" bisogna imparare a gestire i propri percorsi informativi nella maniera più critica e aperta possibile, giocando con l'immaginazione che permette di compiere nuove mosse e di mutare le regole del gioco, inventando azioni capaci di spiazzare gli avversari, proprio come fece l'astuto Ulisse, aggirando con il suo Cavallo di legno le difese che proteggevano le mura troiane dalla sconfitta.

NOTE

1 La bibliografia essenziale è pressoché inesistente. Anche in rete, rivolgendoci a forum e newsgroup, i risultati dell'indagine sono stati scarsi; ci sono studi e ricerche in merito che si focalizzano soprattutto sulla questione tecnica e scientifica, di carattere prevalentemente universitario e sperimentale.

2 Luis J. Prieto introduce per primo il concetto di *pertinenza* in campo semiologico, trattando gli aspetti della fonologia, che viene poi esteso nelle analisi linguistiche. Cfr. Prieto J. Luis, *Pertinenza e pratica*, Feltrinelli, Milano 1976.

3 Questo fatto ricorda il principio di indeterminazione di Heisenberg, secondo il quale la ricerca della precisione si scontra non con un limite tecnologico (l'inefficienza delle strumentazioni) ma con un limite interno alla natura della materia: l'incertezza delle misurazioni, invece che diminuire all'aumentare della precisione, aumenta con essa (I. Prigogine).

4 Questa non è una cosa nuova, come ci ricorda de Saussure: "Dato che una parola

può esprimere idee piuttosto diverse senza che la sua identità sia seriamente compromessa (cfr. '*adottare* una moda' e '*adottare* un bambino', 'il *fiore* del melo' e 'il *fiore* della nobiltà' ecc.). Il meccanismo linguistico ruota tutto intero su identità e differenze, queste non essendo altro che la controparte di quelle", F. De Saussure, *Corso di linguistica generale*, Laterza, Bari 1967, p. 132.

5 V. su questo punto il cap. 5 dedicato alla procedura di indicizzazione.

6 Information Retrieval (IR), recupero delle informazioni. Tutti quei sistemi e operazioni che permettono di ricercare, consultare online, scaricare sul proprio hard disk le informazioni presenti in rete. Si può recuperare le informazioni accedendo ai tre principali modelli di gestione elettronica: data-base (banche dati), file server (raccolta di file in specifici nodi della rete, chiamati server), servizi informativi integrati (unico servizio che offre funzionalità diverse: posta elettronica, accesso a chat, newsgroup, data base e file server).

7 Università e istituti di ricerca tipo Stanford, Waseda, Northwester, MIT.

8 La logica fuzzy è una branchia della matematica inventata nel 1965 dal dr. Lotfi Zadeh. La logica fuzzy permette di esprimere il valore di una variabile con termini appartenenti al linguaggio umano e quindi intende la realtà in modo alternativo: non esistono soltanto il vero o il falso, il caldo e il freddo, ma vi è anche una verità di mezzo, una verità "sfumata" e un livello di temperatura intermedio, il tiepido.

9 Al momento uno dei motori ufficiali esistente che offre un "dialogo" più naturale, in inglese, con l'utente è ASK JEEVES, www.askjeeves.com. Esso pesca le risposte sia nel proprio database sia in quelli di Altavista, Excite, Lycos, WebCrawler e Infoseek.

10 Le metodologie principali per accedere alle informazioni sono: i *Data Base Management System* e il *Full Text Retrieval*. La prima viene utilizzata per il recupero dell'informazione già strutturata, per esempio nel caso in cui è già segmentata in record e field, dove le relazioni all'interno del testo sono stabilite a priori. Si pensi, per esempio alle tabelle che vengono usate per la classificazione di dati aziendali, anagrafici ecc. Il *Full Text Retrieval*, lavorando sul testo non organizzato, comporta operazioni più complesse e articolate, il cui scopo è quello di ritrovare termini e frasi esatte di un testo che è a sua volta inserito all'interno di un vasto data-base, come potrebbe essere un'enciclopedia digitale.

11 Il padre delle attuali procedure di indicizzazione è lo SMART system. La prima versione fu costruita tra il 1968 e il 1970 da Gerard Salton. Per approfondimenti è possibile consultare il materiale e scaricare gratuitamente l'ultima versione di SMART all'indirizzo: ftp://ftp.cs.cornell.edu/pub/smart; il testo di riferimento principale è G. Salton, *The SMART Retrieval System: Experiments in Automatic Document Processing*, Prentice-Hall, New Jersey 1971.

12 In questo caso il "comando" da inserire nel form per trovare tutti i termini che condividono la radice art è: "art*".

13 Questo paragone con alcune correnti dell'arte contemporanea non ha nessuna pretesa estetica, se non quella di esemplificare i concetti.

14 Ogni lingua ha un suo universo carsico di rapporti linguistici e classificazioni sotterranee che, per loro natura, vengono spesso solo "percepite più che comprese", e rimangono il più delle volte inconsapevoli. Whorff definisce queste sottili classificazioni "criptotipi", e cioè "significati sommersi, sottili ed elusivi, non corrispondenti ad alcuna parola reale, ma funzionalmente importanti nella grammatica", come per esempio il fonema del *th* inglese, la cui consolidata presenza nelle particelle dimostrative (*the, this, there, then*) ha creato, a livello inconscio, una "pressione psichica contro l'accettazione del suono dolce *th* in parole nuove o immaginarie che non avessero un significato dimostrativo" (Whorff, *Language, Thought and Reality*, Mit Press, 1956, p. 76). Nei movimenti italiani degli anni Settanta si consolidò un subliminale uso linguistico della K come fattore connotativo della scrittura militante extraparlamentare, il cui significato era dato dal contrassegno fonetico della k (lettera

peraltro inusuale nel contesto linguistico italiano) che faceva assumere a tale fonema precisi compiti semantici e non solo.

15 Nell'immaginario del cyberpunk è innegabile la rivisitazione, per quanto eclettica e originale, del Situazionismo – reinterpretato alla luce dell'esperienza punk e della Xerox Art. A sua volta il movimento del '77, con tutto il suo bagaglio di slogan deturnanti, è tuttora serbatoio inesauribile di idee (saccheggiate dal cannibalismo pubblicitario) e giochi linguistici di irrefrenabile carica surreale: "Lama non l'ama nessuno", coniato dagli Indiani metropolitani, è un esempio di sintesi creativa più folgorante di tanti manifesti politici. Rimandiamo per un approfondimento critico alla lettura di una storia orale del '77, *Una sparatoria tranquilla*, Odradek Edizioni.

16 Un'interessante interpretazione sul filo rosso di continuità dei movimenti, sia politici sia artistici dell'ultimo secolo – da Duchamp all'hackmeeting – è oggetto di studio da parte della sociologa Tatiana Bazzichelli (in "Simultaneità" e in "Neural", 2000; www.strano.net/bazzichelli).

17 Si può assimilare al branching model, ma all'ancora più esplosivo e deflagrante modello della schizoanalisi delineato da Deleuze in *Differenza e ripetizione*. La schizofrenia e gli stati psicotici come esplorazione di un *altrove* linguistico e concettuale altamente creativo, sono anche al centro dell'opera di scienziati quali Stanislav Grof e Gregory Bateson.

18 Per una lettura completa dell'Internazionale situazionista si rimanda alla rassegna completa dei suoi bollettini in *Internazionale Situazionista 1958-69*, Nautilus, Torino 1994.

19 Parafrasando le profetiche parole di Lyotard a proposito del sapere scientifico nel secolo della privatizzazione della ricerca.

20 Ovvero gli accordi tecnico-commerciali per cui si premia economicamente il sito che tramite banner appositi è riuscito a dirottare parte dei suoi visitatori su un altro sito sul quale ha acquistato il prodotto pubblicizzato dai banner.

21 L'Ulisse di Dante nel canto XXVI dell'ultimo viaggio.

22 Atti di censura espliciti che si concludevano nei celebri roghi di libri proibiti.

23 Il loop in cui cade una tale logica di potenza è forse uno dei rischi più inquietanti dell'evoluzione di queste pratiche. Quando infatti la pratica diventa legge (nel caso dei motori, il consolidamento del ranking effettuato in questi termini restrittivi di *doorways* e *keywords pompate*), viene falsato il percorso stesso della ricerca dell'informazione.

24 Stanislav Grof, psichiatra clinico, esperto di terapia psichedelica e inventore della "respirazione olotropica", in *The Cosmic Game* ha esplorato le intuizioni provenienti dagli stati non ordinari di coscienza, giungendo a formulare un quadro filosofico di tipo sistemico in cui si integrano gli sviluppi esperienziali di evoluzione della coscienza con le teorie della complessità (in "Altrove", n. 7, Nautilus, Torino 2000; www.ecn.org/sissc).

25 Non si trattava infatti di una banale formula commerciale come invece è avvenuto in Italia con la finta formula dell'"Internet gratis" che si risolve invece in una fidelizzazione commerciale dell'utenza finale, costretta oltretutto a pagare in termini di perdita non solo della privacy ma anche di risorse economiche per la connettività.

26 www.xs4all.nl

27 www.ecn.org

28 L'idea di Tommaso Tozzi era quella di creare "operazioni subliminali" per espandere "virus informativi" in opposizione al sistema ufficiale della cultura e della comunicazione di massa. "Hacker Art come sviluppo parallelo a ogni tipo di guerriglia clandestina. Hacker Art come definizione di una soglia entro cui l'operazione virtuale e l'operazione reale si scambiano indistintamente i ruoli mantenendo stabile la gestione e lo scopo dell'atto trasgressivo" (in *Opposizioni '80*, Edizioni Amen, Milano 1991, p. 283).

29 Come in McLuhan ne *L'occhio a l'orecchio* avviene uno spostamento dal significato di visibilità a quello di orecchiabilità.

30 In una carrellata di sperimentazioni linguistiche che più hanno plasmato l'ultimo secolo, non può mancare l'autore di *Alice nel paese delle meraviglie*, il quale in ogni sua opera letteraria o matematica ha affrontato il complesso tema dei paradossi logici. Il movimento della controcultura ha sempre reso omaggio a questo onirico scrittore d'avanguardia.

31 Ma è doveroso citare anche il contributo di autori quali Rimbaud, Orwell, A. Huxley, P.K. Dick, Vonnegut, W. Gibson, N. Stephenson, R. Quenau, I. Calvino, G. Perec, Don DeLillo, N. Balestrini, i testi dei CCCP, e lo sperimentalismo musicale degli Einstürzende Neubauten, e DVA... Ma un capitolo a parte di interesse linguistico dovrebbe essere ricoperto dallo studio dello sperimentalismo chimico-psichedelico degli anni Sessanta. Il nome della dietilamide dell'acido lisergico, il cui acronimo è LSD, ha battezzato un'epoca. LSD, psylocina, MDMA, THC, ayahuasca, ibogaina ecc. – insieme a tecniche come il biofeedback o la vasca di isolamento sensoriale – hanno fatto da laboratorio per tutte le sperimentazioni emozionali e psicosomatiche che hanno dato un forte impulso innovativo nel campo conoscitivo e linguistico.

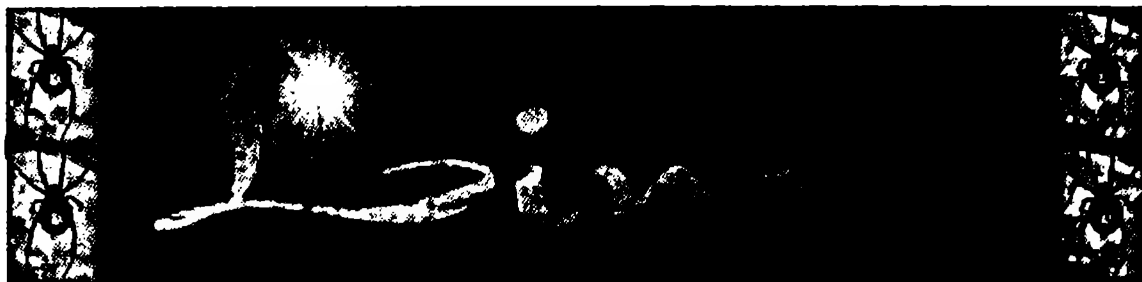
32 La vicenda di etoy www.etoy.com è emblematica ed è stata oggetto di un approfondito articolo in "Neural" (2000). etoy: corporation di sette artisti austriaci, svizzeri e inglesi nata nel 1994. Esponente del "digital hijack", meccanismo per mettere in trappola il navigatore dirottandolo per mezzo di un indirizzamento forzato all'interno del Web. Cfr. Nico Piro (a cura di), *Cyberterrorismo. Come si organizza un rapimento virtuale*, Castelvecchi, Roma 1998.

33 Da citare il Luther Blisset Project, e le esperienze di Tactical Media Crew (www.tmcrew.org) ed ECN (European Counter Network) in generale, i cui contributi ad un approccio linguistico comunicativo nuovo sono da anni sotto gli occhi di tutti.

PARTE TERZA

TECNICHE DI SOPRAVVIVENZA

Peer to peer - condivisione - orizzontalità - sharing - free source - scambio - socializzazione dei saperi



Linux puo' rimpiazzare Windows su un PC o windows NT su un server. Nel mondo ci sono diversi milioni di utenti di Linux che lo usano nei più disparati ambienti. Linux viene utilizzato come sistema operativo sui server principali da molti siti web e ISP. È usato come workstation per il design da molti grafici. Viene utilizzato come principale piattaforma di sviluppo per molti programmatori C e Java nel mondo.





FREENET

Peer to peer - condivisione - orizzontalità - sharing - free source - scambio - socializzazione dei saperi



7

AGENTI SOFTWARE O AGENTI UMANI?

7.1 LIMITI E PROSPETTIVE DEGLI AGENTI INTELLIGENTI

Che cosa sono gli agenti di ricerca? Si tratta di assistenti personali in grado di eseguire per conto dell'utente e in modo praticamente automatico numerose attività connesse all'elaborazione di dati,[1] come, per esempio, la ricerca di una certa pagina web e la creazione di una sua copia sul disco rigido.

Ma in cosa differiscono dai motori di ricerca?

A differenza dei motori, gli agenti sono istruiti da una serie di informazioni che riguardano l'utente e che devono influenzare la ricerca stessa da parte dell'agente. Altra differenza con i motori: gli agenti iniziano a ricercare le informazioni a partire da un percorso prestabilito, così come fanno i motori, ma sono poi liberi di utilizzare gli agganci ipertestuali di questo percorso iniziale, con la possibilità di muoversi per la rete lungo tracciati anche largamente imprevisti.

Un agente di ricerca è il famoso software WebCompass[2] che peraltro utilizza proprio i motori di ricerca come base di partenza per le sue esplorazioni ipertestuali; un altro esempio di agente di ricerca può essere un browser offline come teleport[3] o wget[4] che, opportunamente configurati, consentono lo scaricamento di pagine web selezionate secondo una lista di parole-chiave di partenza.

Ma perché questi agenti si definiscono "intelligenti"? In realtà gli agenti attualmente utilizzati dispongono di funzioni automatiche ma sono privi di intelligenza. In futuro però si prevede che gli agenti di ricerca, anche grazie agli sviluppi dell'intelligenza artificiale saranno in grado di valutare molteplici fattori determinanti nella scelta di una decisione.

Le decisioni che deve prendere un agente intelligente nel reperimento dell'informazione consistono principalmente nel valutare la pertinenza di un'informazione trovata, in relazione alle caratteristiche del profilo personale dell'utente che ha ordinato la ricerca. Da questo segue che quanto maggiore e dettagliata sarà la definizione che un utente è riuscito a dare di sé, tanto maggiore sarà il grado di pertinenza dell'informazione trovata, ovviamente in relazione anche alla specifiche richieste dell'utente stesso.

Questo ci conduce però a uno dei problemi più gravi riguardanti l'uso degli agenti intelligenti: la loro capacità di movimento (o mobilità).

Infatti, se per uno scopo relativamente semplice[5] come la ricerca dell'informazione non è necessaria una particolare interazione con il server che l'agente sta visitando, in altri casi – come quelli legati a operazioni di e-commerce – sorge il problema di far eseguire all'agente dei codici sul server, il quale, sebbene sia in grado di riconoscere alcune configurazioni e funzioni dell'agente relative a criteri di sicurezza (per esempio l'emissione della firma elettronica che identifica l'agente e le sue attività),[6] può decidere proprio per ragioni di sicurezza di non far continuare l'agente nelle sue attività.

Così come il server non si fida a far eseguire il codice all'agente, anche l'utente che istruisce quest'ultimo per l'esecuzione di una ricerca o un'operazione di commercio elettronico potrebbe avere dei problemi di sicurezza. Pensiamo all'esempio di dover acquistare un oggetto in rete e alla relativa necessità di attrezzare l'agente con gli estremi della propria carta di credito: il timore del furto di dati così sensibili a opera di un server appositamente configurato allo scopo è legittimo.[7]

Ammesso e non concesso che i problemi di mobilità degli agenti[8] – connessi soprattutto alla loro capacità di esecuzione di codice sui server incontrati nelle loro esplorazioni – vengano risolti in tempi brevi, sarà attuabile il sogno di creare strumenti configurati al fine di svolgere ricerche in rete, ispirandosi proprio a criteri costruiti *ad personam*: ricerche su misura effettuate valutando se l'informazione trovata o il percorso da intraprendere è non soltanto conforme alla richiesta, ma anche pertinente con le caratteristiche dell'utente. Un agente che sa di compiere una ricerca per una persona di lingua inglese, amante dello sport e in particolare del tennis su erba, desiderosa di ricevere sulla propria scrivania ogni mattina le news inerenti al suo hobby, acquisirà informazioni differenti e sceglierà di muoversi fra link diversi dell'agente configurato da quelli per l'utente di lingua italiana, amante del calcio o del tennis su

terra rossa. Entrambi gli agenti di questi utenti saranno istruiti in base a una griglia informativa di coordinate-base che scandaglia i più comuni motori di ricerca oppure i canali push specializzati in tematiche sportive; ma fin dai primi passi opereranno scelte diverse nel reperimento dati e nella scelta del percorso ipertestuale di ricerca. La rete viene così "radiografata" dagli agenti trasversalmente ai suoi millepiani; gli agenti cercano di fiutare tracce originali e non percorrono unicamente le autostrade prestabilite dei motori.

My Yahoo! è un esempio di agente di ricerca che invece imposta la base di ricerca affiancandovi a un motore e vi rimane addirittura all'interno. Tra le altre cose, My Yahoo! permette di utilizzare un metodo di analisi che classifica i risultati in base alle preferenze impostate dall'utente: una sorta di ranking personalizzato sulle proprie caratteristiche personali.

Gli agenti di ricerca in effetti tendono a interagire – per non dire confondersi – con gli altri principali strumenti di ricerca della rete, mimetizzandosi con i motori di ricerca e con le risorse umane online.

Firefly,[9] per esempio, è un agente con capacità di apprendimento che consente la creazione di un profilo personale, costruito acquisendo informazioni su specifici interessi dell'utente in fatto di divertimento (musica, cinema ecc.); sulla base dei requisiti suggerisce poi una lista di nomi di utenti con gusti simili permettendo a tali utenti di comunicare tra loro.[10]

Looksmart[11] offre un ampio ventaglio di servizi quali l'interazione con redattori online e con comunità di appassionati raggruppati per materie di interesse comune.

Un ruolo di primo piano è sempre più frequentemente quello svolto dagli agenti umani. La pratica dell'intelligenza collettiva nelle sue reti specifiche, definita da De Kerckhove come "intelligenza connettiva", è una risorsa dalle enormi potenzialità che viene intesa come complementare dagli agenti e motori di ricerca, ma è sempre più considerata indispensabile per poter offrire servizi qualificati e personalizzati alle esigenze di ognuno.

I motori di ricerca sono una risorsa immediata per costruire una piattaforma informativa di base su cui lavorare o ripartire per ricerche più approfondite, mentre gli agenti di ricerca – se ben configurati – offrono un ventaglio di possibilità per reperire, almeno in parte un'informazione più specifica e personalizzata. D'altra parte, entrambi questi strumenti risentono dei limiti imposti dai propri automatismi. Vediamo quali.

Innanzitutto, entrambi si trovano a elaborare l'informazione in base a dizionari linguistici e indici di parole-chiave che devono inevitabilmente tenere conto della parzialità perenne delle interpretazioni di espressioni linguistiche. Torniamo all'esempio precedente: abbiamo una persona appassionata di sport e in particolare di calcio, e immaginiamo che l'a-

gente di ricerca trovi per lui un'informazione o un link accompagnato dalla frase "el pibe de oro". Tale espressione è chiaramente una metafora e non contiene nessuna parola chiave utile all'agente di ricerca per individuare pagine web o seguire nuovi link presumibilmente connessi alla tematica del calcio. Eppure agli occhi di un appassionato è indubbio che quel link rimanda a una pagina dedicata all'ex campione del mondo e giocatore del Napoli Maradona! L'agente di ricerca è capace di comprendere solo le parole o le frasi che gli sono state indicate dall'utente e quindi gli sfuggiranno tutte quelle parole pertinenti ma scritte in lingue non definite dall'utente, oppure frasi con ortografie scorrette o appartenenti al linguaggio idiomatico.

In secondo luogo, l'agente è incapace di scegliere autonomamente la base informativa di partenza dalla quale avviare l'esplorazione dell'informazione cercata. Resta indispensabile un'impostazione umana, non solo nella selezione delle parole-chiave che guideranno la ricerca, ma anche nella scelta di particolari portali, motori di ricerca, strumenti push o semplici pagine web la cui esistenza viene segnalata per vie molteplici[12] (visione di telegiornali, consiglio di persone amiche, segnalazione su una mailing-list, ascolto di uno spot alla radio, percezioni subliminali ecc.). In alcuni casi è persino necessario che l'agente software venga affiancato – a volte anche sostituito – da un agente umano: una persona che per le sue relazioni sociali, affinità intellettuali, capacità professionali e abitudini personali sia costantemente aggiornata su ciò che succede rispetto a determinate tematiche, o abbia gli strumenti adeguati – perché no, anche software – per reperire le informazioni richieste.

In poche parole, un cacciatore di informazione.

7.2 CACCIATORI DI INFORMAZIONE

L'agente umano ha una capacità superiore nel reperire l'informazione, perlomeno dal punto di vista intuitivo ed extrarazionale, rispetto a uno software. Facciamo l'esempio concreto di un cacciatore di informazione che debba fornire un'informazione altamente specifica. Comincerà con il cercare e frequentare le mailing-list, i canali IRC e i newsgroup pertinenti, semplicemente sintonizzandosi sulla lunghezza d'onda della comunicazione in atto, o anche ponendo interrogativi che ottengono quasi sempre risposte interessanti e pertinenti, suscitando catene di nuovi interrogativi e risposte... Inoltre chiederà aiuto ai propri amici e forse sfoglierà qualche rivista del settore e ascolterà programmi radio e tv sull'argomento.

Utilizzerà sicuramente anche motori di ricerca e agenti intelligenti, ma avrà una visione a 360 gradi del contesto all'interno del quale si colloca la sua ricerca, acquisendo la capacità, più raffinata rispetto a un software, di distinguere e selezionare una varietà di materiali e link pertinenti e interessanti.

Alla fine, l'agente umano riesce sempre a trovare per commissione quello che cercava per sé. Sono infatti sempre di più le testate giornalistiche, i singoli giornalisti, persino i portali e i motori di ricerca organizzati a directory che si avvalgono della competenza di cacciatori di informazione professionisti, di solito attrezzati con ottime postazioni multimediali ed elevate risorse di connettività, ma anche di strumenti di navigazione molto spartani, e comunque in grado di visualizzare numerosi siti, uno di seguito all'altro o anche simultaneamente.[13]

Rientrano indubbiamente nella categoria dei cacciatori di informazione anche quei servizi commerciali che effettuano ricerche di informazione online su richiesta, oppure che forniscono informazioni dettagliate sulla base di raccolte di link specializzate.

Tutti questi servizi possono avvantaggiarsi tanto di operatori umani quanto di automatismi artificiali di vario tipo, ma è ovvio che saranno sempre i primi a garantire una qualità superiore nell'informazione trovata. A conferma di quanto diciamo invitiamo il lettore a utilizzare Ask Jeeves,[14] un servizio che non fa uso di agenti umani.

Ask Jeeves rappresenta un interessante tentativo di mettere l'utente finale in condizione di porre una domanda molto articolata su un certo argomento senza dover conoscere i linguaggi di interrogazione dei motori di ricerca. D'altro canto, il motore non risulta mai essere particolarmente efficace. Consta di un'interfaccia amichevole che permette di porre un quesito su un argomento adoperando un linguaggio naturale, che viene poi tradotto dal servizio in linguaggio di interrogazione comprensibile per alcuni fra i motori di ricerca principali.[15] Ask Jeeves ha il merito di facilitare l'esistenza a chi non ha tempo e voglia di studiarsi la logica booleana e le regole di ricerca avanzata dei vari motori. Finora si può dire che, al di là del produrre domande su domande per cercare di interpretare l'interrogazione umanamente formulata, non fornisce in ultima battuta risposte molto efficaci. Solo uno sviluppo futuro delle applicazioni della logica fuzzy potrà migliorare le tecniche di ricerca quali quelle utilizzate da Ask Jeeves.[16]

Sicuramente migliore è un servizio recentemente inaugurato sul Web italiano e dal nome "VorreiSapere"[17] che, attraverso una semplice procedura di richiesta via e-mail, fornisce in tempi brevi (entro le 24 ore) risposte mediamente esaurienti, in particolare per sapere se esistono link inerenti alle tematiche richieste. È un esempio evidente di come una redazione umana si riveli più efficace di qualsiasi agente intelligente a tutt'oggi elaborato.[18]

Basati, invece, perlopiù su automatismi di vario genere e su agenti intelligenti sono alcuni servizi specializzati in settori specifici: su rbay[19] si può trovare un elenco di utility inclusi quelli detti "online detective", ovvero servizi specializzati nel recupero di informazioni di carattere personale. Careermart[20] invece è un vero e proprio agente specializzato

nella ricerca di opportunità di lavoro, e svolge egregiamente la sua funzione di analisi di offerte su molti siti americani, anche se per funzionare richiede un "pedaggio" all'utente, ovvero gli impone di compilare un curriculum personale dettagliato.[21]

Nell'ultima categoria dei cacciatori di informazione possiamo includere tutti quei soggetti privati o istituzionali che realizzano directory di link commentati.[22] Ne esistono ormai di vario tipo e sono un valido aiuto a chi è in cerca di informazioni, in quanto permettono di valutare se un certo sito vale la pena di essere visitato o meno, fornendo liste precompilate e commentate di link divise per argomento.

È il caso, per esempio, del Social Science Information Gateway (SOSIG), creato per fornire una selezione di link ritenuti interessanti in materia di scienze sociali, economiche e giuridiche. Per ogni sito indicato SOSIG[23] fornisce una scheda di commento comprendente lingua di origine, abstract, parole-chiave, autore del sito e URL di riferimento.

Tutti i servizi recensiti in questo paragrafo affiancano con successo le risorse umane di analisi ai software di ricerca. D'altra parte, queste risorse umane risultano limitate nella loro capacità di ricerca, in quanto – essendo costituite da singoli individui o istituzioni – inevitabilmente filtrano, cioè selezionano e valutano, le informazioni sulla base di particolari criteri. Sono cioè privi di quella possibilità di critica e libertà di espressione che invece riscontriamo in molte delle comunità presenti in rete, la cui ricchezza sta proprio nella libera condivisione del sapere.

7.3 POTENZIALITÀ E SVILUPPI DEGLI STRUMENTI COLLETTIVI DI CONDIVISIONE DEI SAPERI

La rete è piena di umanità. Chi non si riconosce in questa affermazione non conosce la rete o forse è un software o un replicante, sebbene non se ne sia ancora accorto!

In effetti, bisogna essere freddi come un prodotto tecnologico per non riuscire a scorgere in mezzo a tutto il codice e agli automatismi presenti in Internet, sprazzi di calda umanità. Anche in quei casi in cui l'apparenza tenda a mostrare emozioni tiepide o addirittura negative. Anzi, proprio questa potrebbe essere la controprova più forte del suo "essere" umanità.[24]

È difficile razionalizzare, per esempio, la sensazione di sentirsi parte di una comunità in rete oppure riuscire a decodificare emozioni forti dalle friogeniche e distaccate rappresentazioni del cyberspazio. È difficile spiegare tutto questo a chi non ha mai provato, con pazienza e perseveranza, a entrare in una comunità in rete, e forse è impossibile tentare anche una trasposizione di queste sensazioni in un saggio. Ma cercheremo in ogni caso di sensibilizzare i lettori sulla questione, almeno

per porre l'accento sul senso e l'importanza delle community nel contesto della ricerca dell'informazione in rete.

La socialità in rete[25] non deve essere considerata un'alternativa valida alla mancanza di socialità nella vita reale; è incontrovertibile che essa configura un modo di rapportarsi agli altri, adoperando dinamiche e regole che sono specifiche della rete stessa e che difficilmente sono riproducibili al suo esterno. Infatti è difficile in altri ambiti annullare (o transgenderizzare)[26] il concetto di identità fisica e continuare a mantenere, allo stesso tempo, le stesse possibilità comunicative di chiunque altro.

Tutto ciò incide sulla ricerca dell'informazione in modo sottile e impalpabile, un aspetto che necessita di un approfondimento ulteriore. Per fare ciò dobbiamo prima delineare alcuni tratti dello *stare insieme* in rete. Lo "stare insieme", e il coinvolgimento emozionale che ciò comporta sono uno dei primi obiettivi della rete, tanto che la posta elettronica e i newsgroup, due strumenti eccellenti per comunicare in rete, sono stati fra i primi servizi storicamente attivati e anche oggi annoverati fra quelli che riscuotono il maggior successo.

La necessità di creare nuove reti e sottoreti all'interno di Internet, il cui fine è appunto fare "Reti di reti" (Internetworking), viene avvertita fin da subito e non solo per un "generico" desiderio di socialità ma per il piacere di scambiare, attraverso la socializzazione, i saperi acquisiti e soddisfare il (bi)sogno di espandere le proprie conoscenze.

Le comunità in rete – che siano improvvisati canali chat di tipo IRC, affollatissimi newsgroup, oppure ordinate mailing-list – nascono dalle esigenze più diverse: desiderio di socialità nel senso più classico della parola,[27] necessità di coordinarsi fra persone che sono perseguitate o emarginate per i motivi più disparati,[28] voglia (e/o perdita) di identità e appartenenza a un gruppo ben delineato;[29] ma anche, spesso e volentieri, il piacere di condividere saperi faticosamente conquistati e – viceversa – il bisogno di reperire un'informazione malleabile e altrimenti troppo volatile.

Una comunità in rete è un palcoscenico ideale per condividere/mostrare il proprio know-how, ed è anche un formidabile banco di prova per collaudare il livello di qualità dell'informazione reperita e un efficace (non)luogo dove porre domande e sperare, nella maggior parte dei casi, di ottenere risposte.

Parlando di comunità possiamo certamente pensare ai canali IRC, dove spasmodiche e spassose conversazioni in tempo reale sul proscenio pubblico fanno quasi da sfondo a una miriade di confronti privati dove parole ed emozioni forti trovano spesso sfogo in duelli rusticani su base high-tech.[30] I canali IRC sono anche formidabili strumenti dove si può sperare di reperire un'informazione "al volo"; infatti, basta indovinare il canale giusto[31] e porre la domanda in termini precisi e gentili[32] per ottenere il più delle volte una risposta esauriente.

Quando si dice comunità non si può non pensare alle fluide capacità comunicative dei newsgroup. Detti "forum di discussione di tipo pubblico", i newsgroup rimangono oggetti misteriosi per la maggior parte dei novelli navigatori del cyberspazio in quanto – come nel caso dell'IRC – per poter essere utilizzati richiedono un minimo” di conoscenza tecnica sulla rete. Come già accennato, essi rappresentano la "matrice" di base del cyberspazio, ovvero il network dove l'informazione viene spesso generata. In ogni caso rappresenta la faccia pubblica e viva di Internet: infatti se un indirizzo web è sinonimo di vetrina in Internet, newsgroup significa informazione.

Gli annunci più importanti relativi a nuovi software, le discussioni tecniche e politiche più animate, il materiale informativo più raro, scabroso e censurato, sono rintracciabili senz'ombra di dubbio attraverso la rete interminabile dei newsgroup.[14] Questi, a differenza del patinato Web, sono materia viva; in essi l'informazione può essere postata da un italiano, criticata da un finlandese, manipolata ad arte da un francese e subito sconfessata da un australiano. In questo marasma affascinante e terribile di agitati mari dell'informazione è possibile porre una qualsiasi domanda e trovarne la risposta, con l'unico rischio di incappare in risposte fuorvianti o scorrette, o di perdersi in qualche vortice di discussione fine a se stesso, lasciandosi attrarre da qualche scroscio di rumore di fondo della rete che ci distolga dalla nostra ricerca.

Le mailing-list, a detta di chi scrive, sono invece – fra le strutture classiche di comunità della rete – lo strumento collettivo più efficace per la ricerca dell'informazione. Le mailing-list non sono facilmente localizzabili; ci s'iscrive e si partecipa attraverso la posta elettronica e una propria casella postale: per queste ragioni la "lista" è considerata un modo molto più intimo di partecipare a una comunità in rete. Inoltre, sempre per questa ragione, la lista è considerata con molto più riguardo e offre maggiore affidabilità rispetto a un canale IRC o a un newsgroup.

Gli iscritti a una lista seguono con grande attenzione il dibattito in corso, ma non necessariamente vi partecipano. Infatti, avviene comunemente di trovare mailing-list dove sono iscritte anche centinaia di persone, ma il cui dibattito viene alimentato da poche decine. È il caso di cyber-rights,” – una mailing-list – dove, nonostante una massiccia partecipazione d'iscritti e un gateway[16] con il Web[17] che garantisce una vasta visibilità, si registra un numero relativamente basso di partecipanti attivi al dibattito in corso. La maggior parte degli iscritti e la totalità di coloro che la leggono da Web, un po' per timidezza, un po' per paura di deragliare nell'off topic o di dire cose tecnicamente imprecise, si limitano a seguire il dibattito senza intervenire. Parlando di cyber-rights ci pare comunque doveroso ricordare che questa mailing-list ha fortemente contribuito a far nascere e crescere questo libro, fornendo una grande varietà di informazioni tecniche e spunti di discussione.

Le liste sono un contesto fra i più affidabili dove poter effettuare una domanda; o meglio sono una delle forme di espressione di comunità in rete dove è fortemente probabile ricevere una risposta affidabile ai nostri quesiti. In una lista, infatti, si celebra per così dire un processo di "fissazione dell'informazione".

Le fasi di fissazione dell'informazione sono più o meno le seguenti: l'informazione viene in genere richiesta ai partecipanti della lista da chi ne ha bisogno. A quel punto, accade che se qualcuno è già in possesso dell'informazione richiesta, la fornisca subito con generosità; altrimenti scatta un dibattito interno, ma aperto, sulle modalità migliore per reperirla. Se l'informazione consiste in un'interrogazione, scaturisce il dibattito per la sua rapida risoluzione. Alla fine si trova la soluzione al quesito e viene fornita una risposta soddisfacente che innesca l'ultima fase del processo di fissazione dell'informazione.

In che cosa consiste dunque la fissazione? È un processo di pubblico esame del prodotto elaborato e della sua qualità. Se l'informazione scovata o la soluzione ipotizzata difetta in qualche suo aspetto teorico o tecnico, è assai probabile che venga sottoposta a correzioni e miglioramenti a opera di qualche pignolo partecipante alla lista, fino a quando il risultato non risulterà soddisfacente. I vantaggi nel partecipare a una community in rete possono essere quindi ricapitolati in due punti:

a) elevata capacità di reperimento delle informazioni: le fonti di informazione e di ricerca sono rappresentate dagli stessi partecipanti alla lista, tutti agenti umani che, soprattutto in alcune liste frequentate da esperti, sono giornalisti e tecnici specializzati in vari settori, rilevatori di novità e osservatori accorti e curiosi dell'esistente,[38] e costituiscono – potenzialmente beninteso – il meglio che si possa sperare di trovare;

b) capacità di elaborazione dell'informazione: l'informazione (o domanda) viene analizzata, elaborata e metabolizzata dalla discussione e dall'esame collettivo, fino ad assumere una forma che si può considerare altamente affidabile.

Rispetto ai motori e agli agenti di ricerca, le comunità in rete hanno delle forti potenzialità di sviluppo per quel che riguarda la ricerca di informazioni e l'analisi delle informazioni reperite.

Negli ultimi anni, oltre alle classiche strutture di canali IRC, newsgroup e mailing-list,[39] la rete ha però elaborato altri modi di "stare insieme", di fare comunità, alcuni dei quali hanno assunto inevitabilmente caratteristiche interessanti proprio sotto il profilo del reperimento dell'informazione.

Sempre più spesso si formano in rete comunità che elaborano proprie directory con l'intento di presentare una raccolta di indirizzi com-

mentati con un comune orientamento, oppure specializzati in un certo settore o per una particolare utenza. È l'esempio di Altravista,[40] progetto il cui nome (ludico gioco di parole) richiama alla mente la possibilità di fornire altri punti di vista – visioni alternative – ben diversi da quelli formati dai motori di ricerca straconosciuti, e che risolve l'annoso problema che si trova ad affrontare la comunità di Isole nella rete[41] di catalogare il peraltro incensibile mondo dell'underground digitale.

Ma l'orientamento autenticamente innovativo che sta sempre più prendendo piede tra le community in rete è quello del file-sharing. Promotori del file-sharing sono tutti quei progetti quali Napster, Gnutella, Freenet e molti altri ancora,[42] che, utilizzando protocolli e procedure diverse tra loro, permettono a tutti di raggiungere un risultato simile: la condivisione di file, documenti o addirittura parti intere di disco fisso (o memoria).

Tra quelli elencati Freenet, forse, è il progetto più ambizioso e capace di rivoluzionare l'utilizzo della rete nel prossimo futuro. Nato dalla fervida mente di un giovane studente irlandese, Ian Clarke,[43] Freenet[44] si propone di realizzare una sottorete capace di proteggere l'anonimato e la libertà di espressione dei suoi partecipanti. Una volta scaricato e installato il server freenet su un computer in cui è presente una Java Virtual Machine, è possibile immettere in questa speciale rete documenti e file di qualsiasi natura contrassegnabili da nomi (detti chiavi)[45] oppure ricercare file particolari, purché si conoscano le rispettive chiavi. I file saranno caricati o scaricati senza che sia possibile rintracciare con precisione la locazione degli stessi file o l'indirizzo degli utenti che compiono operazioni sui file. Tutto ciò è elaborato sulla base di un principio filosofico di protezione dell'anonimato, che proprio tramite Freenet trova la sua massima espressione.[46]

L'idea del file-sharing, corollario della filosofia Freenet, consiste nel mettere a disposizione interi settori del disco rigido del computer che svolge funzioni da server Freenet, senza che il gestore del server si renda immediatamente conto[47] di quanto stia succedendo e sia di conseguenza in grado di decidere se accettare o no sessioni di condivisioni di file e documenti. La sola cosa che su Freenet è impossibile fare è cancellare operazioni e file, almeno fino a quando il server è online: in questo modo la censura viene tecnicamente bandita.

Freenet è già funzionante fra una ristretta cerchia di ammiratori, e si ritiene che in futuro possa rappresentare un'alternativa valida ai sistemi di controllo e reperimento dell'informazione in rete. Come Linux, anch'esso partorito da una giovane mente,[48] è un prodotto free, ovvero aperto all'analisi e al contributo della comunità in rete. Come Linux[49] ha attirato su di sé l'attenzione delle corporazioni commerciali, tanto che al momento i fondatori dei due sistemi stanno lavorando anche per conto di holding informatiche molto potenti e indirettamente collegate al co-

losso Microsoft. È la naturale conseguenza di un Mercato che scopre che la rete gira al meglio intorno e dentro le comunità. Comunità che, una volta consolidate attraverso una precisa campagna di fidelizzazione (e di censimento di dati privati) possono rappresentare un'importante fonte di guadagno.

Finora Freenet è riuscita a resistere al richiamo dei facili guadagni provenienti dalle corporazioni commerciali restando ancora uguale a quello che era al momento della sua nascita: una rete libera. Lo stesso non si può affermare per Napster,[50] un altro servizio di scambio di file musicali, nato con lo stesso spirito che ha permesso la costituzione di Freenet, ma ora gestito da una società finanziaria.

Si tratta, come è noto, di un sistema di condivisione di file MP3, con il client Napster gli utenti collegati a un server Napster possono scambiarsi i files MP3 presenti nei loro pc. Di recente Napster è salito alla ribalta perché l'industria discografica ha fatto pressione anche in ambiti giudiziari, affinché venisse chiuso d'autorità. Una chiusura di questo tipo è inutile perché Napster, in quanto servizio, può essere facilmente sostituito da una grande varietà di servizi analoghi. Tant'è che già oggi esistono moltissimi cloni di Napster e di progetti di file-sharing specializzati nella condivisione di file MP3, e alcuni anche di carattere no-profit come il famoso GNUtella[51] che, pur essendo ancora lento rispetto a un network di tipo Napster,[52] offre la non commercializzazione dei dati degli utenti che partecipano allo scambio di file e la possibilità di condividere qualsiasi altro tipo di file.

Ipotesi: e se non fossero solo file di tipo MP3 a essere condivisi attraverso un network apposito, bensì i bookmark delle nostre navigazioni? Proviamo a pensare a una comunità di navigatori del cyberspazio che, accomunati da un interesse, da una passione o un orientamento culturale, e dotati di appositi client connessi ad appositi server, potrebbe efficacemente e in tempo reale abbreviare e ottimizzare le proprie ricerche e navigazioni in rete, scegliendo di (con)dividere le proprie risorse informative e le tracce di navigazioni ipertestuali.

Qualcosa di affine è stato già elaborato nel passato. Era un progetto di *(active) collaborative filtering* denominato "Simon", in cui gli utenti condividevano liste di link ipertestuali a pagine web appositamente commentate. Purtroppo di questo progetto non si trova più traccia in rete; ma ciò non toglie che l'idea di unificare il formato di rappresentazione dei bookmark, in modo tale da poter ottenere liste/indice facilmente scambiabili in rete,[53] rimane un'ipotesi di intelligenza collettiva – e connettiva – che potrebbe concretizzarsi in un futuro prossimo, trasformandosi in un formidabile strumento di ricerca dell'informazione in rete[54] basato sul principio della libera condivisione dei saperi.

NOTE

1 Secondo la definizione data da Leslie Lesnick e Ralph E. Moore nel loro, *Agenti di Ricerca*, edito da McGraw-Hill Libri Italia nel 1997.
2 http://www.quarterdeck.com.
3 http://www.tenmax.com/teleport/pro/home.htm.
4 http://sunsite.auc.dk/wget/index.html.
5 Da un punto di vista delle risorse di esecuzione del codice necessarie.
6 Con la firma elettronica il server può quindi distinguere un agente ritenuto innocuo da uno potenzialmente nocivo.
7 Operazione non di facile realizzazione tecnica, ma sicuramente fattibile.
8 La capacità di trasferire codice di tipo Java da un computer a un altro rappresenterà un'opportunità importante per gli agenti intelligenti del futuro.
9 http://www.ffly.com.
10 La società privata Agents Inc. creatrice di Firefly e originata dal famoso Media Lab del Massachussets intende così testare questa tecnologia per estenderla ad altri campi commerciali, quali i servizi finanziari e il turismo.
11 http://www.looksmart.com.
12 E per loro natura difficilmente registrabili da uno strumento software.
13 Nell'ambito dei cacciatori di informazione va molto di moda per esempio sia il browser lynx, perché molto veloce, ma anche il browser Opera, in quanto capace di visualizzare efficacemente più siti contemporaneamente.
14 www.askjeeves.com.
15 "Can You tell me some news about art and in particular on visual arts?" verrà probabilmente tradotto per Altavista in: "news AND visual AND art*".
16 Molti matematici ritengono che solo una sofisticata ricerca linguistica fondata sulle regole fuzzy potrà discriminare le sfumature del linguaggio naturale (mai semplicemente binario) permettendo di interpretare i frattali linguistici che rendono ancora difficile una gestione unicamente software della ricerca in rete.
17 www.vorreisapere.it.
18 La ricerca nel settore dei motori di ricerca è comunque molto intensa e non è detto che non ci regali qualche sorpresa in un prossimo futuro. Si consiglia allo scopo di tenere d'occhio la directory www.agentbase.com/survey-main.htm.
19 http://rbay.com/spy/spy.htm.
20 www.careermart.com.
21 Si è usato il termine "pedaggio"in quanto è ormai noto che le informazioni personali rappresentano un valore anche economico nell'universo di Internet.
22 Seppure con un proprio carattere originale è da citare il caso dei notiziari spediti via e-mail sempre ricchi di URL utilissime e aggiornate come http://www.zeusnews.com.
23 http://sosig.esrc.bris.ac.uk/.
24 È indubbio, per esempio, che quando si scrive un messaggio in una mailing-list e non si riceve risposta è segno il più delle volte non tanto di indifferenza ma di silenzioso assenso. È però vero anche il contrario: ben difficilmente passano indenni nelle comunità in rete affermazioni prive di qualsiasi senso.
25 Ampiamente analizzata da Luciano Pacagnella, cit.
26 Il concetto di transgender è uno dei cavalli di battaglia di Helena Velena.
27 Inutile nascondersi che la ricerca di emozioni rappresenta anche in rete la spinta prevalente per comunicare.
28 La rete è sempre stato un rifugio per i maledetti, i perseguitati politici, gli intolleratamente diversi.
29 Altrimenti come spiegarsi l'esistenza di newsgroup come alt.alien.vampire.flonk.flonk.flonf?
30 Chi partecipa a una sessione chat con IRC così come con ICQ si espone inevitabilmente, tramite il proprio numero IP, a possibili attacchi e tentativi di intrusione.
31 Si chiamano canali, e sono per la maggior parte pubblicamente elencati e rintraccia-

bili, le decine di migliaia – considerando anche le diverse reti e i diversi server IRC esistenti – di forum a cui poter partecipare o chiedere aiuto.

32 Si ha ovviamente più possibilità di ricevere una risposta utile se la domanda la si effettua in un canale dove si è riconosciuti per qualche ragione, se non altro per la semplice e periodica partecipazione allo stesso.

33 Peraltro si allude proprio a un livello minimo di conoscenza tecnica che consiste nel sapere a quale newsserver connettersi, avendo a disposizione una connessione e un apposito software newsreader.

34 I newsgroup sono anch'essi decine di migliaia e difficilmente censibili data la loro veloce crescita ed evoluzione.

35 Mailing-list ospitata su Isole nella rete ~ http://www.ecn.org/lists/cyber-rights/ e che vede la partecipazione di più di 400 accounts di posta fra tecnici, giornalisti e appassionati della rete.

36 Si intende per gateway uno strumento tecnico che mette in connessione diretta due ambiti di comunicazione di tipo diverso.

37 In passato anche con qualche newsgroup e qualche area BBS.

38 Consideriamo che ogni partecipante alla lista ha le sue fonti di informazione e quindi una lista rappresenta una ragnatela di fonti fra le più disparate, sia analogiche sia digitali. Un network di bookmark.

39 I quali rispettivamente hanno il pregio dell'immediatezza (IRC), del confronto pubblico (newsgroup) e dell'accuratezza (liste).

40 Altravista, www.altravista.org, presenta a tutt'oggi varie categorie d'interesse: Anarchia, Antifascismo, Antimilitarismo, AntiProibizionismo, Associazioni, Centri Sociali, Collettivi, Commercio equo e solidale, Culture giovanili, Cyber Rights, Ecologia, Terra, Editoria, Immigrazione, Internazionalismo, Lavoro e Sindacalismo, Marxismo, Media Activism, Memoria, Miscellaneous, Musica, Riviste, Scuola, Stragi di Stato, Video.

41 www.ecn.org

42 Molti dei quali sono invece definiti dal termine di *(active) collaborative filtering.*

43 L'ideazione di questa che possiamo definire una sottorete è infatti stata la sua tesi di laurea.

44 L'homepage attualmente è su http://freenet.sourceforge.net/ ma è probabile che si trasferisca velocemente su un proprio dominio indipendente.

45 Provate per esempio a cercare ArticoloHackLabFIBFI e troverete un articolo dell'hacklab di Firenze, una dei primi collettivi italiani di hacktivist che si è rapportato con questo strumento.

46 Cosa questa difficilmente evitabile essendo una rete che comunque si appoggia sulla struttura TCP/IP di Internet, anche se riesce a non utilizzare alcuni suoi principi base di funzionamento come per esempio il sistema DNS usato per individuare i nodi della rete, così che al massimo si riesce a individuare l'IP dei computer con cui si ha a che fare.

47 Uno dei tanti sviluppi previsti di Freenet è quello di crittare e quindi rendere sempre indecifrabile il contenuto di questi file condivisi.

48 Il finlandese Linus Torvalds, cfr. nota 5 del capitolo 3.

49 Con le dovute differenze naturalmente visto che il fenomeno Linux può essere considerato senza timore di sbagliare l'evento informatico dell'anno 2000.

50 www.napster.com.

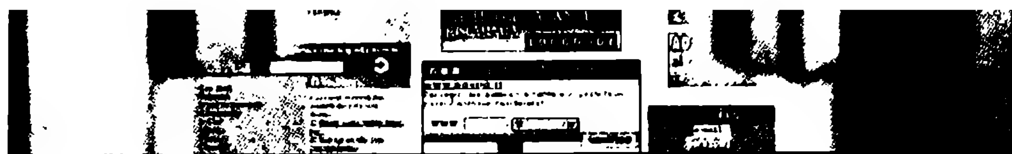
51 La homepage di Gnutella è http://gnutella.wego.com. Va considerato che si tratta di progetti work in progress, la cui natura può radicalmente cambiare nei mesi a venire.

52 La maggiore lentezza è dovuta al fatto che non esiste un server centrale e quindi anche la ricerca si snoda attraverso tutti i pc partecipanti al network GNUtella.

53 A tal proposito si rimanda il lettore alla sperimentazione portata avanti da www.pointera.com.

54 Questa idea è emersa varie volte sia all'interno della lista cyber-rights sia durante alcuni seminari dei vari hackmeeting.

In ambito architettonico cittadino
i portali realizzati da scultori
e maestri scalpellini, in pietra
con tecniche di lavorazione finissime,
abbellivano con l'eleganza degli ornati
le porte di ingresso dei palazzi e chiese
più importanti della città.
Prerogativa di qualsiasi portale era la
rappresentazione, in bella evidenza, dello
stemma nobiliare o religioso, simbolo di potere.



8

FILOSOFIA DEI PORTALI

8.1 TUTTO IN UNA PAROLA

La parola "portale" (portal) compare abbastanza tardi nel vocabolario del Web. Per intenderci, un qualsiasi sito di una certa grandezza dedicato a una serie di temi o ad un argomento specifico era già portale prima che il termine in questione venisse adottato. Si inizia infatti ad usare la parola portale fin dal 1998, anche se i primi grandi network nascono intorno al 1993-1995.[1]

Ma che cosa è un portale? È una "porta" di entrata,[2] uno "snodo" principale della rete, in pratica un punto di riferimento del quale gli utenti si servono per le proprie ricerche o semplici navigazioni; il portale, quindi, può essere considerato, a ragione, una risorsa di ricerca.

Questa specie di porta di accesso che permette di immettersi nel caotico traffico del Word Wide Web è una home page (spesso resa personalizzabile agli utenti) di un sito che opera in rete, svolgendo varie attività ma all'interno di uno specifico settore.

Quello che distingue un portale da un normale sito è la quantità e qualità dei servizi e delle opzioni che è in grado di offrire agli utenti. L'interfaccia di solito molto semplice presenta dei canali tematici o aree dedicate; il motore di ricerca interno è essenziale; possiamo avere un accesso gratuito Internet, una e-mail gratuita; oppure uno spazio web;

o anche la possibilità di inviare messaggi SMS, consultare mappe stradali, elenchi telefonici, meteo, oroscopo, quotazioni di borsa, traduttori, vocabolari; possiamo ricevere newsletter e partecipare a forum, chat, mailing-list; ci possiamo dedicare allo shopping e all'acquisto tramite aste o utilizzare anche i servizi Wap.

Attualmente la parola "portale" è sempre più spesso accostata ad altri termini che cercano di definire più propriamente le caratteristiche e le finalità dei vari portali esistenti: generalista, verticale, orizzontale, tematico, commerciale, trasversale ecc. Ma al di là delle differenze sottili introdotte da queste nuove terminologie, la sostanza non cambia di molto e difficilmente l'utente medio riesce a capire e spiegare la differenza reale tra un portale verticale e trasversale. Forse l'unica grande distinzione che può essere fatta è quella tra portale che contiene più aree e categorie (generalista) e portale che concentra il suo lavoro su un unico tema, il vortale (da vortal, vertical portal).[3]

8.2 L'IMPORTANTE È FIDELIZZARE

Al fine di tenere il passo con la concorrenza, ogni portale deve aggiornare costantemente i propri contenuti, deve essere in grado di attirare a sé nuovi utenti, e infine deve riuscire a mantenere quegli utenti che ha già conquistato. Tale compito, definito "fidelizzazione", consiste nell'invogliare l'utente, mediante le più svariate e pittoresche forme di "attrazione",[4] a navigare stabilmente con fedeltà verso quel determinato sito-portale. Deve venirsi a creare un rapporto "fiduciario" così come accade nel mondo televisivo, quando il telespettatore sceglie il proprio canale per vedere il telegiornale o la partita di calcio. Il parallelo con la televisione potrebbe apparire non completamente pertinente, soprattutto considerando la diversità dei due mezzi di comunicazione, ma in realtà in ambedue i casi il contenuto proposto è sempre scelto cercando di massimizzare l'interesse di grandi bacini di utenza in modo tale da accrescere il numero dei "fedeli" fruitori.

Garantire un notevole numero di utenti che passano attraverso un portale per andare verso altre direzioni, significa in pratica possedere una certa audience,[5] e la possibilità di ottenere consistenti ritorni economici in termini di pubblicità. I banner pubblicitari, ospitati sulle pagine web, avranno dei prezzi assai diversi se i portali che le contengono sono più o meno transitati.

Si deve inoltre far notare che i portali cercano in tutti i modi di immagazzinare i dati personali degli utenti che, come abbiamo evidenziato in precedenza, rappresentano anche un valore economico. La famosa formula che ha preso piede anche in Italia del *Free Internet*, oltre a celare accuratamente il fatto che resta comunque da pagare la connessione telefonica verso Internet, nasconde insidie sibilline. All'atto di

stesura del contratto per il Free Internet, gli utenti vengono costretti a inserire molti dati personali che andranno a costruire profili personalizzati che, al di là delle rassicurazioni legali fornite ritualmente, verranno poi riutilizzati come merce di scambio nell'enorme mercato mondiale dei data-base finalizzati al marketing.

Bisogna inoltre considerare che i portali sono da intendersi come dei "luoghi di ramificazione", dove gli utenti vengono indirizzati in maniera indiretta verso specifiche home page che operano in rete con fini di mero business. Questo indirizzamento è il pedaggio che i naviganti sono costretti a pagare per potersi avvalere, spesso casualmente o per semplice routine, dei servizi che i portali mettono a disposizione. Metaforicamente, il portale può essere inteso come uno specchio telematico per le allodole; un portale che una volta ci offre l'e-mail gratis, poi il servizio SMS, poi ancora il premio in "gettoni d'oro", ma che alla fine ci costringerà, anche solo per sbaglio!, a cliccare un link che ci porterà dove proprio non volevamo andare.

L'orizzonte dei portali non è però così nero come lo abbiamo appena descritto. Esistono delle situazioni, generalmente no-profit, in cui la parola "portale" dev'essere intesa in modo più puro e meno commercialmente contaminato, come nei primi tempi dell'era Internet: volontariato e associazionismo.[6] Questi sono sentimenti e modi di agire che producono benefici agli utenti, che chiedono di essere guidati e che vedono il portale come un utilissimo strumento di ricerca. Va infine detto che è facile prevedere che in futuro le persone collegate a Internet diventeranno più pratiche e scaltre nella navigazione e useranno agevolmente i potenziati motori di ricerca, in modo tale che gli URL saranno raggiunti velocemente e senza intermediari, e il portale quindi sarà sempre meno adoperato come guida.

8.3 IL SITO GRANDE MANGIA QUELLO PICCOLO

È veramente molto difficile disegnare il prossimo scenario relativo al futuro dei portali, ma osservando con attenzione alcuni eventi si possono azzardare delle previsioni. Diciamo che è probabile che ci sarà una progressiva diminuzione dei siti-portale.

La legge della giungla che vige nel presente del WWW ci conferma che solo pochi siti riusciranno a competere nella lotta per la "seduzione" degli utenti, che oltre a crescere vertiginosamente per numero diventano sempre più smaliziati ed esigenti.

Il mantenimento di un portale comporta spese fisse e investimenti finanziari di grosse dimensioni, che i piccoli siti in cerca di fortuna non riescono a sopportare. L'originalità, ovvero l'inventare nuove categorie di operatività, nuovi nomi, slogan, loghi ecc., sebbene sia ancora una carta vincente nelle attività di rete, è sempre più rara, e i margini co-

struiti sull'inventiva si riducono drasticamente.[7] Molti siti, infatti, hanno preferito abbandonare le classifiche dei portali generalisti e hanno cercato invece di costruirsi attorno a categorie del tutto "originali", risultando così primi in assoluto nelle hit; un escamotage concettuale che spesso funziona solo per un periodo limitato.

Molte sono le tattiche adottate per arrivare in cima (tra le quali, per esempio, quella di non praticare link esterni, per non offrire ulteriore visibilità agli altri), ma in definitiva la strategia vincente e più spietata è costituita dal "far fuori" i piccoli siti, conquistandone utenti, contenuti, idee, "teste pensanti", programmatori ecc., a suon di alleanze strategiche, lauti compensi e scalate di Borsa.

Questo millennio nato sotto la "stella" della New Economy, coincide non a caso con l'era dei grandi accordi stipulati dalle maggiori realtà telematiche e multimediali.[8] Il portale più potente risucchia quello più debole, facendo piazza pulita, e diventando quindi il polo di riferimento principe: fusioni, compartecipazioni, matrimoni effettuati dai portali big con i piccoli (al fine di acquisire larghe fette di utenza, garantendosi in questo modo le sostanziose entrate pubblicitarie, e contenuti nuovi, riuscendo così a esaudire le svariate richieste degli utenti, e a mantenere autorevolezza e credibilità).

La situazione attuale, al riguardo, si presenta molto volatile e incerta: i cambiamenti si susseguono in tempi brevi. In fondo ciò che più ci preoccupa è la logica aggressiva e predatrice del più forte. Temiamo che possa accadere quanto è già successo con il mezzo via etere, che si è trasformato da utile strumento per la diffusione di libera comunicazione in sistema commerciale altamente strutturato e gestito da potenti gruppi economici, tutt'altro che aperti al concetto di comunicazione libera.

NOTE

1 Nel 1995, per esempio, viene costituito a Firenze DADA (www.dada.it), uno dei primi Internet Service Provider italiani. Questo "grande sito" si occupava agli inizi di design, architettura, arte, cinema e altri temi culturali. Oggi attraverso Supereva (www.supereva.it) rappresenta uno dei maggiori Portal Community esistenti in Italia.

2 In ambito architettonico i portali realizzati in pietra da scultori e maestri scalpellini, con tecniche di lavorazione finissime abbellivano con l'eleganza degli ornati le porte di ingresso dei palazzi e chiese più importanti della città. Qualsiasi portale mostrava bene in evidenza lo stemma nobiliare o religioso, simbolo di potere.

3 Si veda VerticalNet (www.verticalnet.com), il portale dei portali verticali.

4 News "fresche di giornata", link accattivanti, concorsi a premi per vincere automobili e viaggi esotici ecc.

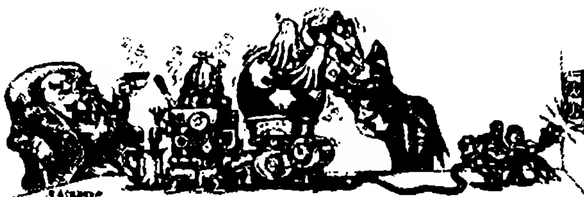
5 Il problema dell'audience in rete è molto caldo; termini tipo *advertising server, impression, audiweb, audinet, click through* ecc. stanno entrando rapidamente nel lin-

guaggio naturale del Web. Vengono continuamente sperimentati sistemi e metodi di analisi che riguardano il controllo e la certificazione delle "visite" ai siti, ma al momento non esistono metodologie ufficiali. Si veda www.abcinteractiveaudits.com, www.mediametrix.com, www.audinet.it, www.doubleclick.net.

6 Ci riferiamo in particolare ai siti *.org.

7 Basta dare un'occhiata all'interfaccia dei principali portali per capire che le differenze strutturali sono minime.

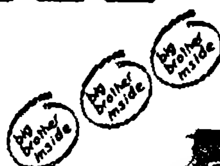
8 Aol e Time Warner, Vodafone-Mannesmann e Vivendi, Seat Pagine gialle e Tin.it, Deutsche Telekom e Club Internet, San Paolo-Imi e Tiscali, tutti accordi che hanno avuto luogo tra gennaio e febbraio del 2000.





Who's watching
your network?



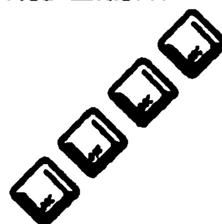




Nel programma del
Chaos Computer Club
in data 1984, anno
dell'assemblea di
fondazione, si legge:
"Esigiamo la realizzazione
del nuovo diritto
umano allo scambio di
informazioni, libero
in tutto il mondo,
non ostacolato e non
controllabile, senza
eccezioni tra tutti gli
esseri umani..."









Se nell'attuale universo della
rete la fruizione assume
caratteri prestabiliti,
con l'assenza dell'aspetto storico
e creativo, in mondo generato dal caos,
la storia entrerebbe in gioco
e l'imprevedibilità svilupperebbe
creatività.







Who's watching your network?



Nel programma del Chaos Computer Club in data 1984, anno dell'assemblea di fondazione, si legge:



9

CAOS VS. BUSINESS?

9.1 FETTE SEMPRE PIÙ PICCOLE DI UNA TORTA SEMPRE PIÙ GRANDE

Quante sono le pagine web? Un miliardo? Cento miliardi? Chi offre di più?[1]

Il 26 giugno 2000, Google ha annunciato di aver censito oltre un miliardo di pagine web, di cui 560 milioni acquisite per intero e 500 milioni segnalate solo in base al loro indirizzo web, ma non ancora analizzate nel dettaglio. Nel frattempo, Lycos amplia l'archivio dei siti catalogati fino a 340 milioni di pagine web, decidendo di utilizzare la tecnologia *fast.*

Non passa neanche un mese che un nuovo studio sulle dimensioni del World Wide Web dichiara che le pagine online superano le più fervide previsioni. Il rapporto di ben 42 cartelle, stilato dalla BrightPlanet,[2] sostiene che in realtà i documenti online sarebbero circa 550 miliardi.[3] Secondo BrightPlanet, Internet sarebbe di qualche ordine di grandezza più estesa di quello che si credeva fino a oggi![4]

A cosa è dovuta questa esplosione della rete? Sicuramente anche all'incremento dei suoi "abitanti". Si pensi che in Cina[5] nel 1997 gli utenti erano 200.000 e nel 2003 saranno circa 35 milioni, in base a una valutazione del "BDA" – Istituto indipendente di ricerca cinese.

Aumentano a vista d'occhio i navigatori, aumentano i webmaster, aumentano i sistemisti e le macchine connesse a Internet: come non prevedere un incremento esponenziale anche delle pagine web presenti in rete?

La rete si sta inoltre differenziando sempre più e i motori di ricerca si dovranno prima o poi scontrare con il problema (irrisolvibile) di catalogare tutta la rete. Inoltre, in breve si troveranno nella situazione in cui sarà impossibile per loro ignorare tutte quelle sottoreti che per motivi commerciali – come nel caso del Wap e Publius – o ideologici – come nel caso di Freenet – si stanno sviluppando ed estendendo indipendentemente dai parametri dettati dal Web.

Se limitiamo il nostro discorso al solo ambito della rete così come è generalmente conosciuta, scopriamo che esistono studi ritenuti molto affidabili che assegnano percentuali di copertura incredibilmente basse ai pur blasonati motori di ricerca. Secondo il dossier di S. Lawrence e C. Lee Giles pubblicato su "Nature",[6] i motori di ricerca prendono in considerazione una minima parte della rete: Altavista il 12%, Northernlight il 16%, Lycos il 2,5%! Inoltre, sempre secondo questo studio, la porzione di Internet indicizzata da questi motori è in gran parte di tipo commerciale.

Internet può quindi essere vista come una grande torta che i motori di ricerca si affannano a ritagliare in tante fettine, sempre più piccole e sempre più infarcite di contenuti commerciali. Eppure, è la fetta non toccata a essere più appetibile: è nella porzione meno nota, quella non censita e meno popolare della rete, quella meno linkata, in quanto estranea ai meccanismi di marketing della rete,[7] che forse, e a volte sicuramente, si cela l'informazione che cerchiamo e che pure non riusciamo a scoprire.

9.2 RANKING: OVVERO COME GLI ULTIMI NON SARANNO MAI I PRIMI

Le classifiche dei risultati di motori di ricerca abbondano di vizi. È indubbio che l'ordinamento dei risultati – il ranking – prodotto dalle interrogazioni dei motori di ricerca sia non completamente attendibile, a causa di numerosi fattori di natura tecnica, ma anche di trucchi finanziari e "culturali". Il ranking, e in particolare il top ranking, è il prodotto di una serie di eventi particolarmente complessi; tuttavia, non è difficile comprendere l'unica logica alla base di questi eventi: esaltare le differenze di risorse che i vari siti hanno a loro disposizione.

È innegabile che nel turbo (o anarco?) capitalismo generato dalla New Economy, non tutti i siti appaiono alla stessa maniera agli occhi di uno spider, dal momento che solo quei siti che dispongono di notevoli capitali e risorse hanno la speranza di arrivare in vetta alle ambite clas-

sifiche. Esaminiamo in dettaglio gli elementi che determinano quali sono i gap fra i siti che determinano la loro presenza o meno nella hit.

È necessario che un sito possa investire in risorse per acquisire il know-how sufficiente a elaborare un codice HTML che si presti efficacemente ad essere acquisito e ben indicizzato dal motore di ricerca. Chi *non* è in grado di ottimizzare un codice agli occhi dello spider del motore difficilmente potrà sperare di entrare nella pole position dei risultati.

I motori di ricerca che hanno anche una struttura di directory ad albero selezionano i siti in base a criteri di bellezza, ricchezza di contenuti, popolarità e perfino, talvolta, in base ad accordi economici.[8] Per esempio, chi *non* ha le competenze per realizzare un bel sito, o *non* ha una redazione[9] capace di produrre in continuazione contenuti di alta qualità, oppure *non* ha risorse economiche da investire per un buon ranking, si trova con le ali tarpate e le ambizioni tradite.

Il ranking di molti motori di ricerca si fonda sul meccanismo di misura della popolarità del sito – popolarità che, lo abbiamo già visto, si misura sulla base del numero di link che in rete puntano a quel sito. Chi *non* ha soldi da investire in banner o in altri accordi di partnership commerciale finalizzati allo scambio o alla collocazione strategica di link in rete, chi *non* suscita scandalo o interesse clamoroso e improvviso in rete, chi *non* ha risorse sufficienti per realizzare un sito su un tema di specifico interesse da propagandare in rete, chi *non* ha le risorse pubblicitarie per catturare attenzione e attirare link sul proprio sito o portale, difficilmente potrà nutrire qualche speranza di visibilità sui motori.

Chi invece decide di investire, anche sotto il profilo economico, per inserirsi nella hit dei risultati ha nel suo carnet di possibilità una strada molto semplice: posizionare appositi banner, quasi trompe l'oeil ben evidenti, vicino alle mascherine che permettono l'immissione della stringa di ricerca sui motori. Infatti, è possibile far apparire il proprio banner in maniera mirata, sulla base di specifiche parole chiave, immesse da un utente per la ricerca, il cui contenuto è simile o vicino a quello reclamizzato dal banner.

È ormai chiaro che essere nelle prime pagine dei motori di ricerca è il risultato di una serie di azioni complesse e tutt'altro che casuali. Viceversa suggeriamo proprio il caso – o meglio il caos – come tool o applicazione concettuale, per tentare di risolvere i problemi del reperimento dell'informazione e della catalogazione dei materiali recuperati. Uno strumento in grado di affrontare la complessità del caos informazionale, capace di farci intraprendere azioni di reperimento e indicizzazione di materiali e contenuti lontani da quelli prepotentemente proposti dal mondo commerciale.

E ciò in quanto non si può affrontare una situazione complessa e non lineare utilizzando un sapere ancora intriso di ideologie e metodologie classiche, tipiche di una scienza ottocentesca e meccanicistica.

L'universo informazionale che emerge dallo stato magmatico e rizomatico della rete è una struttura che assomiglia a un "organismo vivente caratterizzato da un flusso e da un cambiamento continui nel suo metabolismo".[10]

Così ci pare essenziale proporre un diverso approccio di ricerca per esplorare a fondo questo organismo magmatico che è l'universo Internet. All'indicizzazione verticistica e gerarchica dei motori di ricerca forse bisogna sostituire (anche se solo parzialmente) una ricerca caotica e ridondante, e una relativa decodificazione degli eventi che renda conto del mutamento di paradigmi nei saperi contemporanei.

L'ipotesi del "caos" assume dunque un significato provocatorio ma anche antiteocratico, che richiama per l'appunto i paradigmi delle scienze d'avanguardia.

Il caos è fonte di vita; questo è un assunto alla base di importanti ricerche e decisivi studi sia del passato – basti pensare al principio dell'evoluzione di Darwin[11] – sia del presente. Quasi tutta la matematica d'avanguardia, in particolare quella applicata allo studio dei fenomeni naturali, si basa su principi "caotici"; nell'ultimo decennio perfino la chimica di sintesi ha abbandonato le metodologie della chimica ortodossa (e lineare) in favore di un modello di chimica "random",[12] e così l'arte: molti stili e movimenti quali Dada, Fluxus, e il più moderno Dissipazionismo hanno avuto per riferimento atteggiamenti spontanei e casuali, libertari e antidogmatici, volti nella direzione di un'eclettica e dissenziente interdisciplinarietà.

Persino il business prospera su ciò che è generato casualmente dagli esseri viventi: ogni strategia commerciale o di marketing prende spunto da ciò che spesso è generato per caso e poi sviluppato in contesti antistorici, scientifici e politici. In questa prospettiva, anche i motori di ricerca, e soprattutto i portali, non sono altro che appropriazioni e rimaneggiamenti di antiche abitudini e consuetudini dei navigatori della prim'ora, che in maniera pseudocasuale costruivano e distribuivano ai propri amici elenchi di bookmark costruiti con passione e perseveranza in anni di navigazione libertaria nel cyberspazio. Purtroppo, una volta che il "principio del caos" viene raccolto e adottato dal business, viene strutturato, limitato e sfruttato a scopi commerciali. I motori di ricerca e i portali (intendendo chiaramente chi li gestisce) in quanto prodotti di business, non possono non seguire questa strada: nascono avvalendosi di tecnologie sviluppate in ambiti di ricerca, come Internet, spesso frequentati da soggetti quanto meno stravaganti – il "padre" di Fidonet è un anarchico gay – i quali si preoccupano di mantenere la libera diffusione dei loro "prodotti". Dopodiché, utilizzando le intuizioni creative di navigatori esperti e "avanguardisti", si trasformano in prodotti commerciali il cui scopo principale è quello di produrre capitale "tagliando la rete a fette".

Nel momento in cui si avventa il profitto avviene il giro di vite: la generazione spontanea, comunitaria, caotica e vitale lascia il passo al mortifero irrigidimento imposto dallo sfruttamento commerciale, simbolo di un omologato mondo economico spregiudicato e a senso unico.

Come far fronte a questo stato di cose? La nostra risposta è semplice: forse attraverso una pratica sperimentale che abbia fondamento nel concetto di casualità.

9.3 STRUMENTI DI RICERCA CAOTICI COME ANTIDOTI ALL'INFINITO E AI MECCANISMI DI CENSURA

La prima proposta è la seguente: promuoviamo lo studio di sistemi che permettano una scelta casuale della porzione di rete da indicizzare. In altre parole, vogliamo promuovere l'utilizzo di meccanismi che siano orientati al reperimento di informazione in Internet partendo da indirizzi casuali, ma possibili, della rete stessa. In tal senso, un'ipotesi potrebbe essere quella di indicizzare una porzione della rete partendo da IP numerici oppure da nomi simbolici possibili.

Qualora abbracciata fideisticamente questa diventa una proposta dalla forte carica utopistica e dalla improbabile utilità pratica. Tuttavia, se viene presa in considerazione come strategia da affiancare alle attuali politiche di reperimento dell'informazione in rete dei motori di ricerca può avere una sua validità. Si auspica quindi che i motori di ricerca prendano in considerazione la possibilità di utilizzare, anche se soltanto in parte, tale strategia, in modo tale da generare indirizzi simbolici possibili della rete come:

www.parola_possibile_in_italiano.it

oppure

www.word_in_english.org.

e integrare l'indirizzo dell'eventuale sito, trovato per caso, con i relativi link all'interno delle loro basi di conoscenza.

Lo stesso potrebbe verificarsi ipotizzando l'utilizzo di indirizzi numerici IP casuali fra quelli potenzialmente destinati a ospitare server web. Una volta "azzeccato" un indirizzo IP numerico o indirizzo simbolico alfabetico, il motore di ricerca si troverebbe quindi nella rara condizione di poter esplorare e indicizzare una porzione di rete che altrimenti non sarebbe mai stata trovata perché poco conosciuta.

Una pratica quindi di accesso casuale al materiale informativo da reperire in rete che non risolverebbe sicuramente il problema di indicizzare "tutta" la rete, ma che di certo contribuirebbe a rendere più "democratico" il processo di indicizzazione dei siti, minimizzando la possibilità di penalizzare o escludere tutte quelle comunità di rete che, per un motivo o per l'altro, non riescono a essere facilmente visibili.

9.4 LE RAGIONI DEL CAOS CONTRO QUELLE DEL BUSINESS NELLA DEFINIZIONE DELLE CLASSIFICHE DEI RISULTATI DI RICERCA

La proposta avanzata nel precedente paragrafo, pur se teoricamente interessante, si rivela problematica nella sua realizzazione tecnica e presumibilmente poco produttiva. Risulta invece più semplice la progettazione di un'interfaccia che risolva il problema legato alla scelta del criterio da impiegare nella presentazione dei risultati di ricerca.

La soluzione che proponiamo, eticamente più intrigante, sebbene ancora non necessariamente la più efficace è questa: "Quale criterio / nessun criterio".

In altre parole, proponiamo la realizzazione di un ranking caotico.

Ricordiamo che le liste di risultati proposte dai motori di ricerca non sempre e non necessariamente sono costruite privilegiando la qualità dei contenuti dei siti in relazione alle richieste. L'abbiamo già visto: il criterio di ordinamento gerarchico è verticale, è funzione (anche) di pressioni economiche e commerciali. Di fatto, maggiori sono le disponibilità economiche di un sito, maggiori saranno le possibilità che quel sito conquisti i primi posti nelle liste dei risultati.

Al contrario, un ranking che non prenda in considerazione nessun criterio di ordinamento gerarchico, scardinerà proprio il concetto di gerarchia verticale; tra i risultati ottenuti si potrà scorgere, fin dalle prime posizioni, il sito amatoriale accanto a quello professionale, quello sconosciuto ancor prima di quello iperpropagandato, quello graficamente improbabile, ma accessibile, che precede quello esteticamente impeccabile.

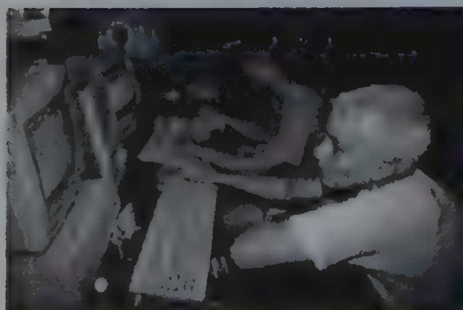
Realizzare un progetto di ranking caotico potrebbe essere relativamente semplice potendo disporre di un metamotore di ricerca che:

1) rigira l'interrogazione fatta a una selezione di motori di ricerca;[13]
2) ottimizza i risultati trovati, escludendo le copie di quei risultati che puntano a uno stesso indirizzo;
3) rimescola l'ordine dei risultati proposti dai motori e li riordina casualmente.

Un'interfaccia neppure eccessivamente sofisticata[14] quindi, che però consentirebbe di avere una classifica di risultati che, proprio perché libera da condizionamenti, potrebbe regalare diverse s/gradite sorprese. Ricchezza – varietà – bellezza dell'incertezza emergono così dalla trama caotica di imprevedibili fenomeni spontanei e instabili. In un mondo generato dal caso "la storia contribuirebbe a giocare le sue carte, il futuro sarebbe incerto e dall'imprevedibilità si svilupperebbe creatività".[15]

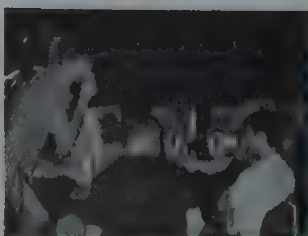
NOTE

1 Il meccanismo del mettere all'asta i bisogni di ricerca dell'informazione non è affatto una provocazione, ma un metodo di lavoro di alcuni motori di ricerca.
2 http://www.brightplanet.com/
3 Sono numerosi gli studi in rete sull'enorme crescita del numero di pagine presenti sul Web. Alcuni esempi: oltre al rapporto di BrightPlanet – http://www.completeplanet.com/Tutorials/DeepWeb/index.asp; Web Bigger Than We Think Yahoo! News http://dailynews.yahoo.com/h/ap/20000727/tc/deep_web_3.html; Study finds Web bigger than we think (CNET) – http://news.cnet.com/news/0-1005-200-2356979.html?tag=st.ne.1002.thed.ni; The Register – http://www.nandotimes.com/technology/story/0,1643,500232662-500338191-501933993-0,00.html.
4 Brightplanet con il suo nuovo motore denominato LexiBot offre la possibilità di effettuare una ricerca su "tutta" la rete censita, a patto che si sia disposti a sborsare 90 dollari e aspettare fino a 90 minuti per una risposta.
5 Esempio significativo, se si pensa che lo sviluppo di Internet è estremamente difficile in questa regione, così come in tutti quegli ambiti dove un apparato politico verticistico è incompatibile con la presenza di un sistema di comunicazione tendente alla decentralizzazione dell'informazione e quindi, di fatto, anche del potere.
6 Vol. 400, 8 luglio 1999, www.nature.com.
7 Quella ovvero che è al di fuori della logica del "cliccka e ti pago", lontana dalle logiche di partnership commerciale e di scambio dei banner.
8 Godado (http://www.godado.it) mette all'asta le posizioni nelle pagine dei risultati. È una specie di partnership evolutasi sulle regole dei motori di ricerca. Chi offre di più compare per primo. La messa all'asta del posizionamento all'interno dell'elenco dei risultati è quantificabile mediamente in 150 lire per keyword, per ogni visita di navigatore che viene sul sito dell'offerente proveniendo da Godado. La scelta delle parole chiave e la serietà/pertinenza dei siti in relazione alle stesse viene controllata da Godado stesso. Godado utilizza anche un altro meccanismo (simmetricamente opposto al precedente, ma ugualmente funzionale) ben consapevole dell'importanza di essere linkati. Godado paga circa 40 lire per ogni mascherina di ricerca di Godado inserita in una pagina web. Altri esempi di motori a pagamento sono DOG PILE, AH-HA e SearchBoss.
9 La "caccia alle teste", ovvero a produttori di contenuti è uno degli aspetti più eclatanti della New Economy: i beni immateriali, saperi – know-how – brain worker, non sono mai stati ricercati in modo così esplicito.
10 F. Capra, *La rete della vita*, RCS Libri, Milano 1998, p. 202. L'opera del fisico teorico consiste in una divulgazione "per tutti" dei concetti più ostici della teoria del caos.
11 Reinterpretato nelle teorie sistemiche in *La rete della vita* di F. Capra.
12 L'attuale chimica combinatoria, in inglese CombiChem, si fonda proprio su sintesi casuali, non lineari e simultanee, vedi Mirella Castigli, *L'altrove della chimica di sintesi*, in "Altrove", n. 6, 1999.
13 Un elenco scelto fra quelli che presumibilmente coprono la maggiorparte della rete.
14 Il mondo della programmazione Free è ormai ricco di prodotti software che coprono ogni esigenza immaginabile; in ambiente gnome-linux esiste anche il metamotore reperibile su http://gwsearch.netpedia.net/ e si presume che possa essere quindi di relativamente facile realizzazione anche un'interfaccia del genere.
15 I. Prigogine e I. Stengers *La nuova alleanza. Metamorfosi della scienza*, Einaudi, Torino 1979. Secondo il fisico matematico inglese Roger Penrose (in *La nuova mente dell'imperatore*, Rizzoli, Milano 1989, p. 22) nella fisica classica l'universo appariva come un immenso automa, sottomesso a leggi deterministiche e reversibili in cui non era affatto facile riconoscere una delle proprietà fondamentali del pensiero: la creatività. Secondo lo studioso, solo introducendo instabilità e caos è possibile rompere la teocrazia dogmatica del determinismo e ridare dignità scientifica alla dimensione creativa. Per non prestare cieca "fedeltà a una razionalità che lascia l'uomo solo in un mondo muto e stupido" (ivi, p. 30).



Mr










APPENDICE MANUALISTICA

In questa appendice abbiamo scelto di provare sul campo una selezione di motori, mailing-list e indici (italiani e internazionali) testandoli con una query molto precisa: la richiesta di una vecchia versione di software per trasmettere e ricevere fax sotto Windows.

Prima di accompagnare il lettore in questo itinerario sperimentale di ricerca pratica dell'informazione, anticipiamo le principali conclusioni derivanti da questo test.

1. Le mailing-list – comunità di rete costituite da esseri umani – si confermano forse come la migliore risorsa per il reperimento dell'informazione in Internet.
2. Per poter ottenere dei risultati utili dai motori di ricerca tradizionali è necessario conoscere le modalità di interrogazione avanzata degli stessi e impegnarsi a farlo per immaginare la stringa di query presumibilmente più efficace per la ricerca in questione.
3. I motori di ricerca tradizionali si trovano in difficoltà di fronte alle sempre più numerose pagine dinamiche presenti in rete che essendo generate in real-time su richieste specifiche, sfuggono alla indicizzazione automatica dei motori e all'opera di ricognizione degli spider in rete. È un problema che può essere risolto, o a livello di gestione del server che seleziona le risposte da dare a seconda del client che le

richiede, oppure realizzando una breve presentazione statica che riassuma i contenuti e le parole-chiave più significative della versione dinamica.
4. Il lavoro redazionale, sebbene selettivo e quindi censorio, risulta comunque in alcuni casi utile per poter reperire l'informazione dai portali in tempi ragionevoli.
5. Resta aperto un problema enorme: come valutare l'informazione trovata e decidere in merito alla sua utilità e affidabilità. Per esempio, se si sta cercando un software (come nel caso che segue) bisognerà anche sapere se il prodotto trovato è affidabile oppure è così mal disegnato che una volta installato manderà in crash una serie di altri programmi o funzionalità del pc che lo ospita, o addirittura se attiverà un virus maligno incluso al suo interno. Qui stiamo parlando di situazioni non particolarmente frequenti, ma che devono comunque essere prese in considerazione. In generale, consigliamo di testare l'informazione trovata (sia essa un testo o un codice binario) con l'arma del buon senso. Anche sotto questo profilo la risorsa della mailing-list appare la più utile, in quanto il risultato trovato viene sottoposto all'esame della comunità tra cui sarà sempre possibile trovare qualcuno capace di argomentare in senso sensato e utile.

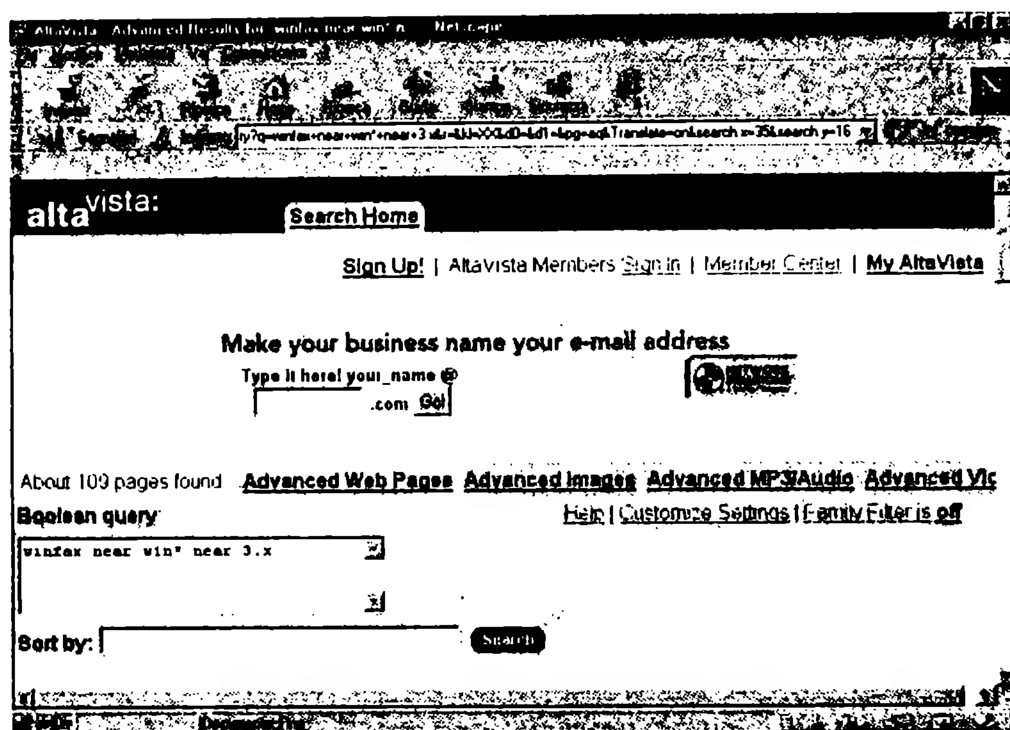
Il primo passo di questo test è consistito nello spedire la seguente mail alla mailing-list cyber-rights [http://www.ecn.org/lists/cyber-rights]:

"ciao all!
un amico, una persona disabile che mi dà occasionalmente mano sul portale dell'accessibilità http://www.ecn.org/xs2web giusto ieri aggiornato mi ha chiesto dove poter trovare una versione di winfax per la vecchia versione di windows 3.1 e/o 3.11
ora mi metto a cercarlo in rete ma c'è qualche anima gentile che mi darebbe la dritta giusta?
magari esiste qualche archivio specializzato in risorse software per hardware povero... magari...
alla prossima!"

Appena spedita la mail ci spostiamo su Altavista cercando prima nella versione di interrogazione standard la stringa "winfax 3.1 e 3.11". Non ottenendo risultati utili al primo tentativo proviamo la possibilità offerta da Altavista delle "related searches", ma anche in questo caso la lista offerta immettendo "winfax download" appare troppo ricca di risultati, pur se risulta leggermente più efficace di quella ottenuta digitando "free winfax download". Decidiamo comunque di utilizzare le proprietà avanzate.

Proviamo anche la funzione di ricerca fra i risultati ottenuti con "free

winfax download" aggiungendoci old e sperimentando anche le chiavi "win or windows 3.x".



Individuiamo un'altra buona chiave, ovvero "WinFax PRO 4.0", che ci sembra essere l'ultima versione per win 3.x

Troviamo alcune pagine con contenuti non corrispondenti a quelli presentati nell'indice: segno che vengono indicizzate di rado... forse abbiamo trovato quello che cercavamo, ma dobbiamo settare il browser in maniera ottimale per poter vedere la pagina correttamente... niente da fare...

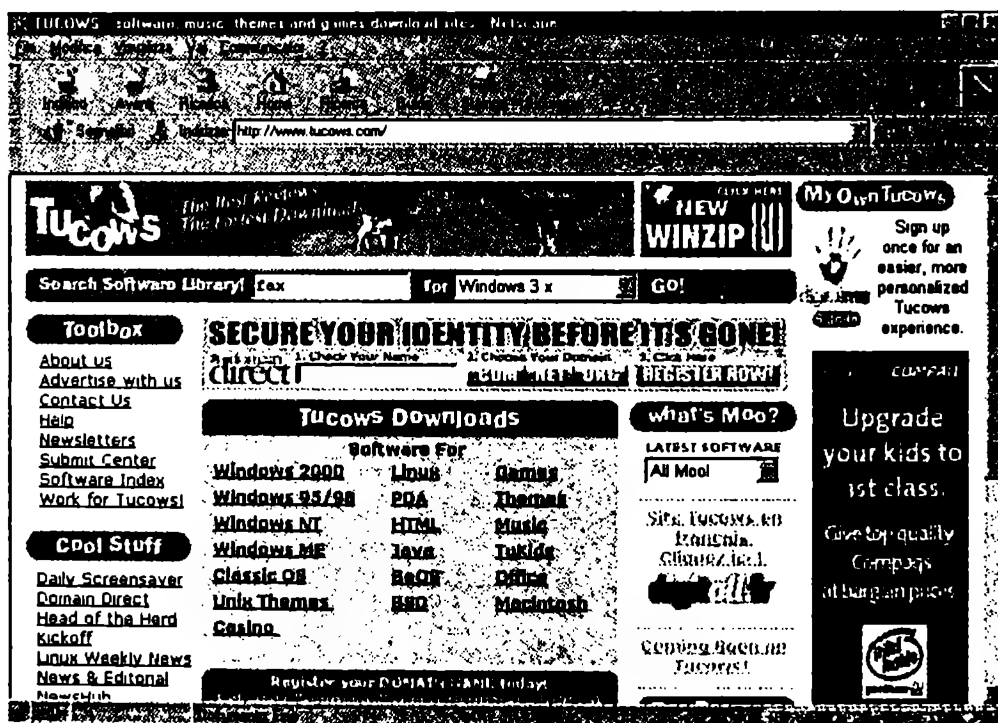
Cambiamo strategia e ci accorgiamo di aver formulato male anche la richiesta nella lista; in effetti, il nostro amico ci chiedeva winfax perché lo conosce e possiede una macchina win 3.x da poter usare, ma potrebbe essergli utile anche un prodotto diverso! Facciamo altri tentativi; in uno di questi troviamo persone a introdurre la stringa "symantec + fax +win* +3.x +download +free", ma non troviamo niente nella mezz'ora che avevamo deciso di investire su Altavista per la ricerca.

Ci spostiamo allora su una directory specializzata sull'argomento software come www.tucows.com e... troviamo subito quello che ci serve.

Dopo poche ore anche dalla lista cominciano ad arrivare messaggi di risposta con indicazioni utili e precise

From: sgnurf@hotmail.com

To: ferry.byte@ecn.org
Subject: [cyber~rights] chi domanda trova ? ;-)
Se ti interessa ti posso uppare QuickLinkII per Windows311 & Dos che mi funzionava bene e occupa 1,2Mb...
Fammi sapere
Non so se l'hai già trovato, anyway segnalo i seguenti archivi per Win3.x:
http://classic.tucows.com/win3xsoft.html
http://www.winsite.com/win3/
http://www.zdnet.com/swlib/hotfiles/w31uti.html
http://ftp.vse.cz/simtel.net/win3/index-win3.html
(non garantisco che tutti i link siano esatti, ogni mese 'sti stronzi riorganizzano i siti...)



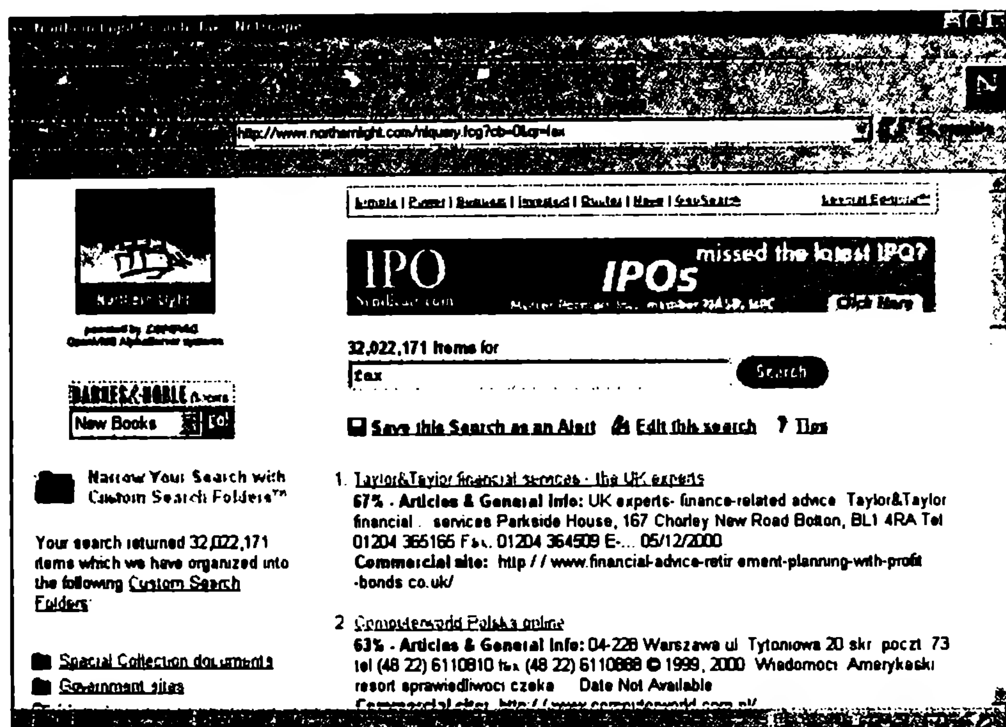
Utilizziamo anche Northern Light, considerato da molti il motore di ricerca che indicizza la porzione più grande della rete, e vediamo cosa troviamo.

Proviamo con una chiave semplice: "fax". Norther Light propone i risultati ottenuti dividendoli fra directory di tematiche specifiche. Ci avventuriamo su "computer software industry" e "all others".

Proviamo anche la power research e aggiungiamo chiavi alla ricerca: "+fax +download +3.x +win +windows +free".

In effetti, muovendosi tra directory consigliate e con la precedente

stringa di chiavi, frutto del lavoro fatto su Altavista, individuiamo velocemente diversi siti con il software che ricercavamo.



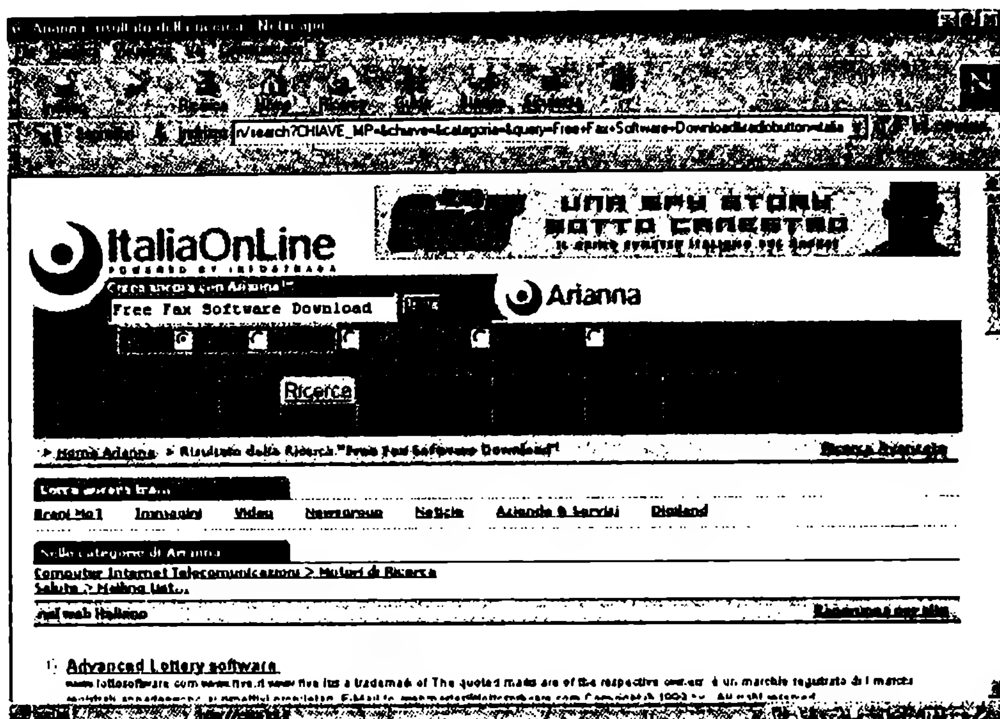
Testiamo infine anche un blasonato motore di ricerca italiano di tipo tradizionale: Arianna.it

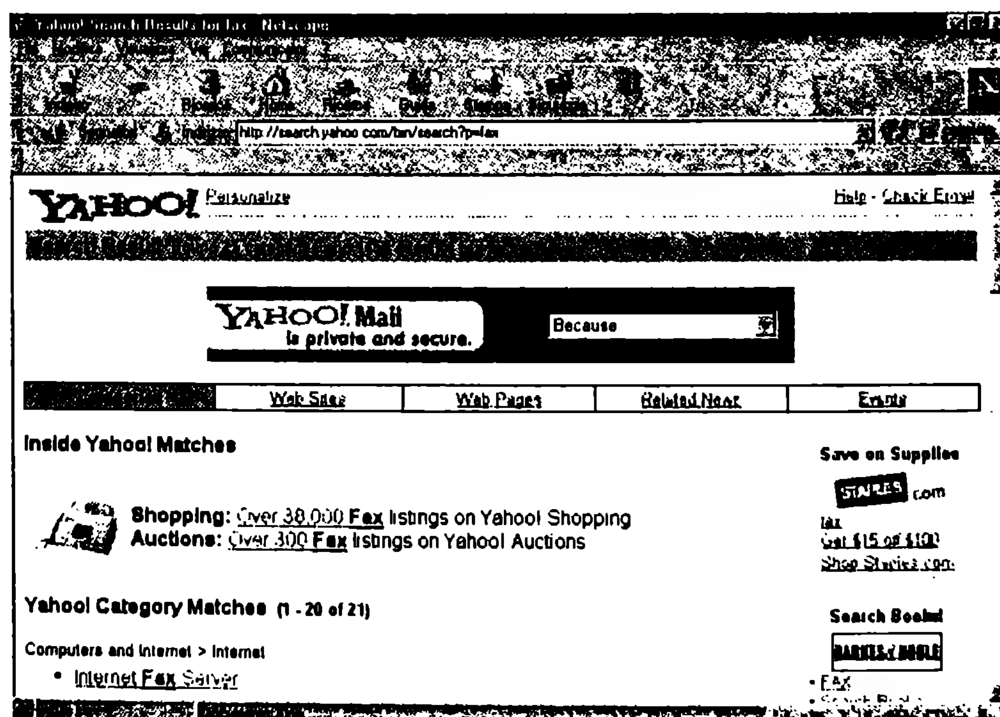
Scegliamo la nuova sezione internazionale, perché "a naso" ci sembra quella più adatta, e digitiamo semplicemente "fax". Subito l'occhio cade sulla directory intitolata "software" e su quella "programs".

In effetti questo sistema, molto simile a quello utilizzato da Northernlight, che propone directory di link su argomenti più specifici sembra funzionare, e ben presto troviamo un elenco di siti utili.

Ovviamente, non possiamo fare a meno di testare anche la directory di rilevanza internazionale per eccellenza: www.yahoo.com.

Digitiamo "fax" e di nuovo ci viene in aiuto un elenco di directory all'interno delle quali però si trovano elenchi di URL poco utili alla nostra ricerca. Azzardiamo un'altra stringa di chiavi: "free fax software download" e qualche altra combinazione di chiavi, ma non siamo fortunati e non troviamo niente per mezz'ora. Proviamo a consultare yahoo nella sua forma classica di directory ad albero, ma anche in questo caso la fortuna non ci assiste e non scopriamo niente di utile in tempi ragionevoli.





Concludiamo la sperimentazione con l'italianissimo indice Virgilio, il quale ha fatto la scelta di mantenere poche directory e pochi URL all'in-

terno del suo data-base: la maggior parte sono indirizzi a loro volta portali di materie specifiche che riescono velocemente a soddisfare i bisogni della nostra ricerca.

BIBLIOGRAFIA

AA.VV., *I motori di ricerca*, Mondadori – I miti dell'informatica, Milano 1999.
Un manualetto di informazioni pratiche su come rapportarsi in maniera utile ai principali motori di ricerca. Un tascabile necessario al neofita, ma anche utile a chi già conosce la materia.

Michel Autier, Pierre Lévy, *Gli alberi di conoscenze*, Feltrinelli, Milano 2000.
"Gli alberi di conoscenze" è uno strumento informatico, ideato dai due autori francesi nel 1992, che consente di mettere in pratica la filosofia della condivisione dei saperi. In questo libro vengono esposti i principi teorici, i meccanismi e le applicazioni di questo sistema di apprendimento e sviluppo della conoscenza. Con un linguaggio agile viene spiegato questo metodo/dispositivo che, con spirito collettivo, aperto e cooperativo, è adottato in ambito educativo, nelle imprese e nelle comunità in genere.

Paul K. Feyerabend, *Contro il metodo. Abbozzo di una teoria anarchica della conoscenza*, Feltrinelli, Milano 1979.
Un pamphlet stimolante e irriverente, oltreché un trattato di epistemologia dadaista, che mostra l'inconsistenza delle ipotesi scientifiche non fondate su fantasia, creatività, astuzia e inventività. Solo un radicale mutamento di paradigmi all'interno della società nel suo complesso, facendo leva su principi antidogmatici, può favorire lo sviluppo di imprese intellettuali libertarie, pluralistiche e animate da un autentico spirito umanitario. L'anarchismo teorico come passepartout per aprirci al progresso scientifico e culturale.

Carlo Formenti, *Incantati dalla rete*, Raffaello Cortina Editore, Milano 2000.
Ricognizione, pressoché esauriente, sul territorio che riguarda Internet: gli immaginari, le aspettative, i timori e le speranze. Seppur con un approccio sostanzialmente neognostico, viene esaminata in modo preciso la svolta epocale che questo mezzo di comunicazione ha prodotto, ricostruendo con accuratezza buona parte delle posizioni filosofiche, antropologiche, tecnologiche, religiose e politiche dei principali protagonisti di questa rivoluzione digitale: cyberteologi, chierici, hacker, anarcocapitalisti.

Matthew Fuller, *ATM*, ShaKe Editions, Milano-London 2000.
L'autore, che fa parte di I/O/D, collettivo londinese impegnato nell'espansione e nella ridefinizione critica del concetto di interfaccia, in questo romanzo breve sviluppa un'estetica del molteplice e, sotto il profilo linguistico, tratta la scrittura come "produzione di contesti e processi materiali reali". Particolarmente interessante è il suo esplorare le macchine linguistiche – "ovvero quei sistemi che producono scrittura che produce la scrittura" – dal punto di vista semiotico.

Leslie L. Lesnick, Ralph E. Moore, *Agenti di ricerca*, McGraw-Hill, Milano 1997.
Mostra le fasi di creazione di un agente (software) "tuttofare" ma propone anche l'analisi in dettaglio di tutti gli aspetti riguardanti la tecnologia degli agenti di ricerca, confrontando le loro capacità di esplorazione della rete con quella dei motori di ricerca tradizionali.

Nico Piro (a cura di), *Cyberterrorismo. Come si organizza un rapimento virtuale*, Castelvecchi, Roma 1998.
Racconta la storia del Digital Hijack (dirottamento telematico); operazione artistica, ideata dagli *etoy*, che indaga criticamente il "linguaggio" dei motori di ricerca, evidenziandone le storture e le potenzialità. Un ritratto a tutto tondo della famosa corporation di artisti, conosciuti anche per la discussa "etoy-s war".

Richard Sennet, *L'uomo flessibile*, Feltrinelli, Milano 2000.
Disamina dello studioso inglese sul fenomeno della mobilità e flessibilità del lavoro, in relazione ai processi di produzione del nuovo capitalismo che sono strettamente collegati alle implementazioni tecnologiche. Una radiografia del mondo del lavoro che cambia radicalmente, sottolineando i rischi e pericoli sulla vita personale.

Varela J. Francisco, Thompson Evan, Eleanor Rosch, *La via di mezzo della conoscenza*, Feltrinelli, Milano 1992.
Con un'acuta riflessione sui fondamenti delle scienze cognitive, prende avvio una profonda discussione sull'idea di verità e di mondo. Un confronto tra tradizione occidentale e orientale che indaga il rapporto tra la conoscenza e l'esperienza, il pensiero e l'esistenza. Un allargamento dei confini della consapevolezza e dello studio sulla mente e sul corpo.

SITOGRAFIA

strano.net/chaos
il sito realizzato dagli autori di questo libro nel quale poter trovare tutto il materiale documentario di riferimento alle tematiche trattate e gli aggiornamenti relativi.

www.tucows.com
portale per il reperimento di software; comprende anche la maggior parte del software descritto in questo libro.

www.motoridiricerca.it
il portale italiano più completo ed efficace sul tema motori di ricerca.

www.ecn.org/loa/terminal/GAIA
il progetto GAIA è nato per risolvere un problema frequente nelle reti a organizzazione gerarchica: la dipendenza da chi detiene un livello di accesso privilegiato alla struttura piramidale.

www.ecn.org
portale delle situazioni di movimento in Italia.

freenet.sourceforge.net
Freenet consiste in una rete di macchine che, pur "appoggiandosi" a Internet, adottano un proprio protocollo di comunicazione che permette loro di ricercare, immettere e scambiarsi informazione senza che sia possibile censurare alcunché o rintracciare gli utenti collegati.

www.decoder.it
la storica rivista dell'underground digitale italiano finalmente sotto forma di sito.

www.northernlight.com
NorthernLight è considerato da molti il motore di ricerca che indicizza la porzione più grande della rete.

www.hackmeeting.org
sito che ospita le edizioni italiane degli hackmeeting, ritrovi socio-tecnici durante i quali, spesso e volentieri, emergono novità anche in tema di ricerca e accessibilità dell'informazione.

altavista.com
a tutt'oggi forse il motore di ricerca più popolare e allo stesso tempo fra i più utili.

arianna.it
il motore di ricerca italiano.

www.google.com
A oggi il motore di ricerca più veloce ed efficace.

www.vorreisapere.it
help online italiano tramite il quale si può inviare una richiesta ad un operatore in carne e ossa che effettua la ricerca, ricevendone poi il risultato per e-mail.

www.delorie.com/web/ses.cgi
Search Engine Simulator.

www.ecn.org/xs2web
portale sull'accessibilità xs2web (Access To Web).

www.w3.org/WAI
Web Accessibility Initiative del World Wide Web Consortium.

www.cast.org/bobby
severo strumento di valutazione dell'accessibilità di una pagina web.

www.tmcrew.org
Tactical Media Crew.

www.ffly.com
agente con capacità di apprendimento.

www.ecn.org/lists/cyber-rights
mailing-list incentrata sui diritti della comunicazione.

www.altravista.org
indice di situazioni "altre".

www.strano.net
gruppo di lavoro sulla comunicazione.

alt.alien.vampire.flonk.flonk.flonf
newsgroup di fondamentale importanza per le sorti dell'umanità.

gnutella.wego.com
sistema non centralizzato per il file-sharing di documenti di qualsiasi tipo.

www.yahoo.com
indice di rilevanza internazionale.

www.virgilio.it
indice di indirizzi italiani continuamente aggiornato e particolarmente efficace in alcune sezioni come, per esempio, quella dell'educational.

www.supereva.it
esempio di Portal Community esistente in Italia.

www.verticalnet.com
il portale dei portali verticali.

www.windweaver.com/searchtools.htm
Search the Net. Your Top Search Resources. A cura di Tracy Marks.

www.notess.com/search/
Search Engine Showdown. The Users'Guide to Web Searching. A cura di Greg R. Notess.

www.unn.ac.uk/features.htm
Web search tool features. A cura di Ian Winship.

www.hamline.edu/library/bush/handouts/slahandout.html
Tips on Popular Search Engines. A cura di Karen Campbell.

sunsite.berkeley.edu/Help/searchdetails.html
Internet Search Tool Details. A cura dell'Università di Berkeley.

magi.com/~mmelick/it96jan.htm
Beyond Surfing: Tools and Techniques for Searching the Web. A cura di Kathleen Webster & Kathryn Paul.

daphne.palomar.edu/TGSEARCH/
How to Search the Web. A Guide To Search Tools. A cura di Terry A.Gray.

www.lib.utexas.edu/Pubs/guides/internetpubs/researchin.html
Internet Resources: Research Strategies. A cura della Texas University.

strano.net/copydown
DOWNload a copy, upLOAD an idea. Una strategia di decentralizzazione delle produzioni artistiche in controffensiva alla nuova legge sul copyright; una TAZ per le produzioni free e no-copyright.

www.searchenginewatch.internet.com
portale di osservazione sui motori di ricerca.

lcweb.loc.gov/global/internet/history.html
la storia di Internet e del www.

www.gmarketing.com
Guerrilla Marketing: strategie di marketing intelligentemente aggressive.

```
_____________________________________________________
_____|    |______
_____| ALT.INTERNET.SEARCH-ENGINES |_______
_____|  CHARTER & FAQ  |______
_____|_________________________________|_______
```

Search Engines
– http://www.altavista.com/
– http://www.alltheweb.com/

– http://www.directhit.com/
– http://www.excite.com/
– http://search.go.com/
– http://www.google.com/
– http://www.goto.com/
– http://www.hotbot.com/
– http://www.lycos.com/
– http://www.northernlight.com/
– http://raging.com/

Image Research
– http://www.ditto.com

Portals directories
– http://www.dmoz.com/
– http://www.looksmart.com/
– http://www.snap.com/
– http://www.yahoo.com/

Submission URLs
– http://searchenginebase.com/sbsumissions.html

Cloaking Tutorial + FAQ
– http://fantomaster.com/fafaqcloak1.html
– http://www.spiderhunter.com/

Keyword Research
– http://fantomaster.com/fasmbres03.html#voyeur

Meta Tags
– The Definitive Resource: http://vancouver-webpages.com/META/

Search Engine Newsletters
– Actu Moteurs (in French): http://www.abondance.com/
– Google Friends Newsletter: http://www.google.com/
– Pay Per Click Search Engines Update: http://PayPerClickSearchEngines.com
– Search Engine Guide: http://www.searchengineguide.com/

Search Engine Optimization Newsletters
– fantomNews: http://fantomaster.com/fantomnews.html
– RankWrite: http://www.rankwrite.com/
– Search Engine News: http://www.searchengine-news.com
– Search Engine Optimization and User Interface:
http://www.cre8pc.com/seui.html
– Search Engine Quarterly: http://www.searchengineworld.com/
– Search Engine Watch: http://www.searchenginewatch.com/
– The Spider Report: http://spider-food.net/



SHAKE EDIZIONI UNDERGROUND



PER ORDINI DIRETTI E INFORMAZIONI

Contatti postali: ShaKe, viale Bligny, 42 – 20136 Milano – tel./fax 02/58317306
Contatti elettronici: http://www.decoder.it e www.shake.it
e-mail: info@shake.it

CATALOGO

CYBERPUNKLINE

NO COPYRIGHT Raf Valvola Scelsi. Per un nuovo diritto positivo dell'uomo. Un testo importante sulla questione della trasmissione del sapere nella società postfordista. Dal software al problema generale, pp. 304, E. 11,88 • **CYBERPUNK** Antologia di testi politici a cura di Raf Valvola Scelsi. 20ª edizione. Il ritorno ruggente della stampa underground, un libro importante che ha disegnato i contorni degli anni novanta, analizzando negli aspetti politic il caso letterario del decennio scorso: il Cyberpunk, pp. 224, E. 10,33 • **SNOW CRASH Seconda edizione riveduta** Neal Stephenson. Primo romanzo edito dalla ShaKe, vero e proprio tecno-giallo dell'età dell'informazione. Premiato come miglior libro di fantascienza ne 1994 con l'Arthur C. Clarke Award, pp. 416, E. 14,46 • **GIRO DI VITE CONTRO GLI HACKER** Bruce Sterling. Testo fondamentale per chi desidera conoscere le dinamiche de cyberspazio, pp. 254, E. 11,88 • **MINDPLAYERS** Pat Cadigan. È il primo romanzo *psycocyber* dell'unica scrittrice riconosciuta dal movimento letterario cyberpunk, pp. 240, E. 11,88 • **HACKERS: eroi della rivoluzione del computer** Steven Levy. È il libro definitivo della storia degli hacker dal 1958 al 1983, pp. 416, E. 17,04 • **AA. VV. STRANI ATTRATTORI - Antologia di fantascienza radicale** a cura di Rudy Rucker, P.L. Wilson (più conosciuto come Hakim Bey) e R. Anton Wilson, ed. it. a cura di M. Tavosanis e F. Gadducci. Una bomba letteraria a base di sesso biomeccanico, droghe in quantità, post-anarchismo militante e tecno-profezie per il nostro caotico futuro. Racconti di J.G. Ballard, W.S. Burroughs, P. Di Filippo, W. Gibson, R. Kadrey, J. Kohenline, T. Metzger, R. Rucker, L. Shiner, B. Sterling e altri, pp. 304, E. 15,49 • **METROFAGA** di Richard Kadrey. L'eccezionale primo romanzo di Richard Kadrey è duro, mai cinico, appassionatamente politico e solleva questioni importanti, pp. 224, E. 12,91 • **L'ERA DEL DIAMANTE** Neal Stephenson. Un romanzo maturo e avvincente, ricco di potenti e inedite immagini del futuro, ambientato nella Cina digitale mandarina, pp. 432, E. 18,08 • **RIBELLIONE NELLA SILICON VALLEY: conflitto e rifiuto del lavoro nel postfordismo** a cura di Raf Valvola Scelsi. Raccolta del meglio degli scritti e della grafica della straordinaria rivista californiana "Processed World" che si occupa da ormai 15 anni dei problemi dei lavoratori del "terziario avanzato", pp. 192, E. 11,88 • **SINTETIZZATORI UMANI** Pat Cadigan. Secondo strepitoso romanzo della "regina del cyberpunk", pp. 400, E. 14,46 • **PERMUTATION CITY** Greg Egan. In una città virtuale le lotte tra ricchi e poveri per la conquista dell'immortalità. Un capolavoro di un grande matematico, pp. 280, E. 14,46 • **RETROFUTURO Antologia di racconti** di Vittorio Curtoni. Questa antologia raccoglie il meglio della sua produzione di racconti brevi, incastonati nella cornice di una personalissima storia della fantascienza ricostruita sul filo della memoria, pp. 256, E. 12,91 • **ZODIAC Un giallo ecoterrorista.** Neal Stephenson. Un vero e proprio thriller con tonalità hardboiled, per un maestro della post fantascienza: un libro veramente divertente, pp.

256, E. 14,46 • **NELL'ANNO DELLA SIGNORA** di Carlo Formenti. In una Lombardia del futuro, il potere è conteso tra congreghe neoreligiose incentrate sul culto della Signora. Un romanzo ben scritto e avvincente, pp. 224, E. 12,91 • **MILANO 2019: LINEA DI CONFINE** di Roberto Perego. Un romanzo velocissimo dallo stile ficcante, uno sguardo su uno degli scenari prossimi futuri, pp. 288, E. 12,91 • **FOLLI** di Pat Cadigan. In un mondo di succhia-cervello e rapinatori di corpi, non puoi dare nulla per scontato. Neanche la tua identità. Per la seconda volta, dopo *Mindplayers*, Pat Cadigan vince il prestigioso A.C. Clarke Award come miglior romanzo di fantascienza, pp. 224, E. 12,91 • **IL CUORE ESPLOSO** di John Shirley. Sussultano i nervi alla lettura dei racconti di questo libro, tenetevi a portata di mano dei sedativi altrimenti il cuore potrebbe veramente esplodervi. John Shirley porta i suoi lettori a dubitare della propria salute mentale, aggredisce i preconcetti fondendo allucinazioni hi-tech con immagini viscerali provenienti dai luoghi più oscuri della sua mente drogata. pp. 192, E. 11,88. • **23 La storia dell'hacker Karl Koch** di Hans-Christian Schmid e Michael Gutmann. Tra le pagine meno note della storia dell'hacking, quelle relative alla vicenda di Karl Koch, Pengo e un pugno di aderenti al Chaos Computer Club sono sicuramente tra le più importanti. pp. 152, E. 11,88 • **IL CYBORG Saggio sull'uomo artificiale** di Antonio Caronia. A distanza di quindici anni esce una versione riveduta e ampliata in cui l'autore propone una nuova riflessione sull'"uomo artificiale". pp. 136, E. 9,30.

UNDERGROUND

PASTO NUDO a fumetti Professor Bad Trip. Il classico sulle droghe illustrato dal grande cerimoniere acido. Introduzione di Fernanda Pivano, pp. 80, E. 10,33 • **L'OCCHIO NELLA PIRAMIDE** Primo volume della trilogia *GLI ILLUMINATI* R. Shea e R.A. Wilson. Questo romanzo si snoda tra magistrali fantasie paranoiche e un'intrigante analisi della realtà "occulta", pp. 304, E. 11,88 • **VERE ALLUCINAZIONI** Terence McKenna. Il più grande studioso di psichedelia contemporaneo, in un romanzo verità, in un viaggio nella foresta amazzonica alla ricerca del più potente allucinogeno, pp. 248, E. 11,88 • **CENTRI SOCIALI: geografie del desiderio**, a cura di Consorzio Aaster, Centro sociale Leoncavallo, Centro sociale Cox 18, Primo Moroni. Le statistiche, il passato, il presente e il futuro dei centri sociali milanesi, pp. 256 (con inserto di 64 pp. di foto), E. 12,91 • **TRAVELLER E RAVER: racconti orali dei nomadi della nuova era** R. Lowe e W. Shaw. Le vere storie di 30 "nomadi della nuova era". Parlano i protagonisti di un movimento che solo in GB conta mezzo milione di aderenti, tra vecchi hippy e nuovi raver, gente che si sposta continuamente a bordo di furgoni, camion e auto scassate, pp. 224, con immagini a colori, E. 12,91 • **ESTATE D'AMORE E DI RIVOLTA Con i Beatles nella Summer of Love** di Derek Taylor. Un'entusiasmante viaggio nella Summer of Love con i Beatles, Rolling Stones, Doors, Jefferson Airplane, Fugs, Ginsberg, Leary, pp. 304, E. 14,46 • **COSTRETTI A SANGUINARE Romanzo sul punk 1977-84** Marco Philopat. Un libro sulla vita di una generazione che ancora oggi porta sulla propria pelle le ferite e le cicatrici di un drammatico passaggio d'epoca, pp. 224, E. 12,91 • **LA MELA D'ORO** Secondo volume della trilogia *GLI ILLUMINATI* R. Shea e R.A. Wilson. Nuove fantastiche avventure aspettano i lettori "intrippati" nella paranoica visione del mondo di quel folle di Robert Anton, pp. 224, E. 12,91 • **HIPHOP A LOS ANGELES Rap e rivolta sociale** Brian Cross. Il testo documenta con interviste, foto e saggi le radici storiche del rap e la varietà delle sue reti culturali, pp. 272, E. 14,46• **HELL'S ANGELS** Hunter S. Thompson. Il mito degli Hell's Angels, le violenze vere o presunte, lo scontro con il potere, le reazioni isteriche della stampa. Il tutto scritto da un temerario giornalista divenuto uno di loro, pp. 272, E. 14,46 • **ATTI INSENSATI DI BELLEZZA. Hippy, punk, squatter, raver, eco-azione diretta: culture di resistenza in Inghilterra** George McKay. Questo testo tratteggia un resoconto vivido e dall'interno (l'autore è stato punk, anarchico, occupante) della rete controculturale e dei movimenti che fino a oggi sono rimasti, per larga parte, non raccontati. Una nuova interpretazione delle zone temporaneamente autonome di Hakim Bey, pp.

248, E. 14,46.• **IL LEVIATANO** Terzo volume della trilogia *GLI ILLUMINATI* di R. Shea e R.A. Wilson. Conclusione dell'intricata saga, misteri e segreti mai rilevati si dispiegheranno al lettore. pp. 240, E. 12,91. • **ALICE È IL DIAVOLO Storia di una radio sovversiva** Collettivo A/traverso, a cura di Gomma e Bifo. La vera storia dell"radio libera" che cambio il volto della comunicazione via etere, con cd allegato che ripropone le voci e i suoni delle manifestazioni bolognesi del marzo 1977, pp 196 + cd E. 30,00.

I PIRATINI

T.A.Z. Hakim Bey. Dalle comunità dei pirati di Captain Mission alle comunità telematiche Cyberpunk. L'elogio del nomadismo psichico in un saggio che è già un classico delle controculture, pp. 176, 80 foto, E. 10,33 • **VIA RADIO** Hakim Bey. Dallo stesso autore di T.A.Z., una raccolta di brevi saggi sull'immediatismo, una strategia di accesso al piacere, "la Tv è per l'immaginazione quello che il virus è per il Dna: la fine", pp. 64, E. 4,13 • **UTOPIE PIRATA: corsari mori e rinnegati europei**, Peter Lamborn Wilson, esperto di *zone temporaneamente autonome*, mette a fuoco le caratteristiche dell'organizzazione pirata: corsari, sufi, pederasti, irresistibili donne, schiavi, avventurieri, ribelli irlandesi, ebrei eretici, spie britanniche ed eroi popolari radicali... la popolazione di queste ormai mitiche *comunità insurrezionali*, pp. 160, E. 9,30 • **MILLENNIUM Dalle TAZ alla rivoluzione** Hakim Bey. Dopo Taz la riflessione si sposta sull'attualità di una prospettiva rivoluzionaria. Il centro è adesso sulla possibilità dell'eliminazione della stato a seguito dell'alleanza tra società e tradizioni religiose ereticali, pp. 96, E. 9,30.

BLACKPROMETHEUS

CON OGNI MEZZO NECESSARIO Malcolm X – I discorsi e le interviste dell'ultimo anno d vita. Uno strumento indispensabile per la conoscenza delle lotte radicali nere, pp. 224, E 12,91 • **SENZA ILLUSIONI** a cura di Bruno Cartosio. Antologia sui neri americani dalle Black Panther alla rivolta di Los Angeles. Contributi dei migliori saggisti afroamericani, pp 272, E. 10,33 • **PANTERE NERE** Paolo Bertella Farnetti. Il libro sull'organizzazione politica che più ha acceso l'immaginazione di varie generazioni di giovani rivoluzionari e ribelli: i Partito delle Pantere nere, pp. 288, 100 foto, E. 12,91 • **IL POPOLO DEL BLUES** Amiri Baraka (LeRoy Jones). In questo famoso testo Baraka ricostruisce la vicenda del popolo afroamericano, dalla fase della schiavitù alla "cittadinanza" politica, attraverso la chiave di lettura offerta dalla storia della musica blues e jazz, pp. 208, E. 12,91. • **LA FABBRICA DEI NEGRI** Gil Scott-Heron. Sincopatico romanzo ambientato negli anni Settanta in una università nel sud degli Stati uniti tra gli studenti neri, pp.208 E. 10,32. • **BLACK MAMA - LA VEDOVA NERA** Iceberg Slim. Durissimo e disperato romanzo considerato negli Stati Uniti uno de più importanti classici della letteratura afroamericana, pp. 256, E. 14,00 • **KING SUCKERMAN** George P. Pelecanos. È un inno al cinema blaxploitation, alla musica funky e soul, al basket e in generale alla cultura pop urbana, pp. 256, E. 14,00.

CORPIRADICALI

LA SADICA PERFETTA Terence Sellers. Il libro rappresenta la *summa* filosofica della sua riflessione sul disciplinamento dei corpi e delle menti... sul bondage, il feticismo, il rituale, la morte. Completano il testo una decina di tavole di Genesis P-Orridge, fondatore degli Psychic TV e una riflessione di Helena Velena sulla funzione psicoterapeutica della sadica "professionale", pp. 208, E. 14,46 • **AD ALTO RISCHIO Antologia di scritti proibiti.** Con scritti di Kathy Acker, William S. Burroughs, Pat Califia, Dennis Cooper, Terence Sellers, Gary Indiana e molti altri. Con questo volume si entra "ad alto rischio", ci si immerge nelle zone proibite del sesso e dei comportamenti trasgressivi e politici, pp. 272, E. 14,46 • **NEW**

NOIR John Shirley. Una serie di racconti noir al limite dello splatterpunk, comici e talvolta irresistibili. Un piccolo cult, pp. 128, E. 9,30 • **AVANT POP** A cura di Larry McCaffery. Con scritti di Samuel Delany, Kathy Acker, Mark Leyner, Euridice, William T. Wollmann e altri, pp. 224, E. 12,91 • **SANGUE SESSUALE** di Mark Amerika. Amerika, uno dei massimi scrittori dell'underground americano, è il portavoce dell'avant-pop, la corrente letteraria che mescola generi e immaginari diversi: dalla cultura beatnik al cyberpunk, dal noir al sesso estremo, pp. 192, E. 12,91 • **IL FETICISMO DELLA SECONDA PELLE Antologia di "Skin Two"**. Una raccolta dei testi migliori della rivista inglese, che non mancheranno di affascinare non solo i lettori che ne hanno abbastanza del "sesso noioso", ma anche tutti coloro che hanno intenzione di esplorare territori proibiti e limiti spesso invalicabili, pp. 192, E. 12,91 • **DON CHISCIOTTE** di Kathy Acker. Questo romanzo, Acker, punta l'attenzione sull'identità della donna usando un modello fortemente maschile: Don Chisciotte. Com'è nel suo stile decostruisce il soggetto femminile e uno dei miti maggiormente presenti nella rappresentazione della donna: l'amore romantico con le sue convenzioni monogamiche ed eterosessuali, pp. 240, E. 14,46. • **F/32 La magni*fica* assassina** di Eurudice. Un romanzo selvaggio, pericoloso, disturbante, metafisico ed eccessivo. pp. 208, E. 12,91 • **SENTO ODOR DI ESTHER WILLIAMS** di Mark Leyner. Dello stesso autore di Mio cugino il gastroenterologo, il libro che fonda la corrente Avant Pop. pp. 136, E. 9,30.

RE/SEARCH RI/CERCA

RE/SEARCH Edizione italiana W.S. BURROUGHS B. GYSIN. Nuova edizione aggiornata, dopo la morte dell'unico scrittore veramente geniale dei nostri tempi, pp. 272, E. 12,91. • **RE/SEARCH Edizione italiana J.G. Ballard**. Antologia del grande anticipatore del futuro prossimo venturo. Contiene 29 saggi originali inediti di non-fiction, edizione totalmente rifatta rispetto all'originale, pp. 272, E. 10,33 • **RE/SEARCH Edizione italiana MEDUSE CYBORG Antologia di donne arrabbiate** Terzo volume dell'edizione italiana di Re/Search, comprende una serie di lunghe interviste a "donne arrabbiate" che rappresentano diverse esperienze e una comune tensione verso un femminismo moderno e radicale, pp. 352, E. 16,53 • **RE/SEARCH Edizione italiana MANUALE DI CULTURA INDUSTRIALE** a cura di Paolo Bandera. Il testo è una guida alla filosofia e all'attività di un gruppo di artisti *devianti* che si è distinto per aver avuto in anticipo l'intuizione del crollo della civiltà industriale, pp. 288, E. 14,46. • **DAVID CRONENBERG** di Serge Grünberg. Cronenberg, proveniente dal movimento underground degli anni Settanta, è un cineasta ossessionato dal mostruoso e dalla letteratura, tanto da fare del problema della mutazione il filo rosso della propria ricerca estetica. Con un saggio inedito su *eXistenZ*, pp. 152, E. 10,33, illustrato. • **GLI OCCHI VERDI** di Marguerite Duras. In questo volume, pubblicato in Francia per i "Cahiers du cinéma", e fino ad oggi inedito in Italia, Marguerite Duras, esprime spietati e contraddittori giudizi sulla società utilizzando come metafora il cinema, pp. 104, E. 10,33. • **GILLES DELEUZE POP-FILOSOFO** di Massimiliano Guareschi. Questo volume si propone di fornire un'imtroduzione, una via di accesso al pensiero deleuziano, esplicitamente pensato per i non specialisti di cose filosofiche, pp. 168, E. 10,33.

UNIVERSALE SHAKE

L'AUTUNNO DEGLI STATI UNITI - Neoliberismo e declino sociale da Reagan a Clinton di Bruno Cartosio. La drammatica polarizzazione sociale e i 15 anni di *downsizing*: questi i temi principali del libro, che propone un cammino pluriennale di analisi critiche, ostinate e puntuali sulla scena economico politica e sociale degli Stati Uniti, pp. 240, E. 12,91 • **YOU'LL NEVER WALK ALONE Il mito del tifo inglese** di Rocco De Biasi. In un volume chiaro e gradevole, vengono analizzate la trasformazione del calcio inglese e delle sue tifoserie e la creazione del mostro sociale "hooligan", pp. 224, E. 12,91 • **DA LIVERPOOL A SAN**

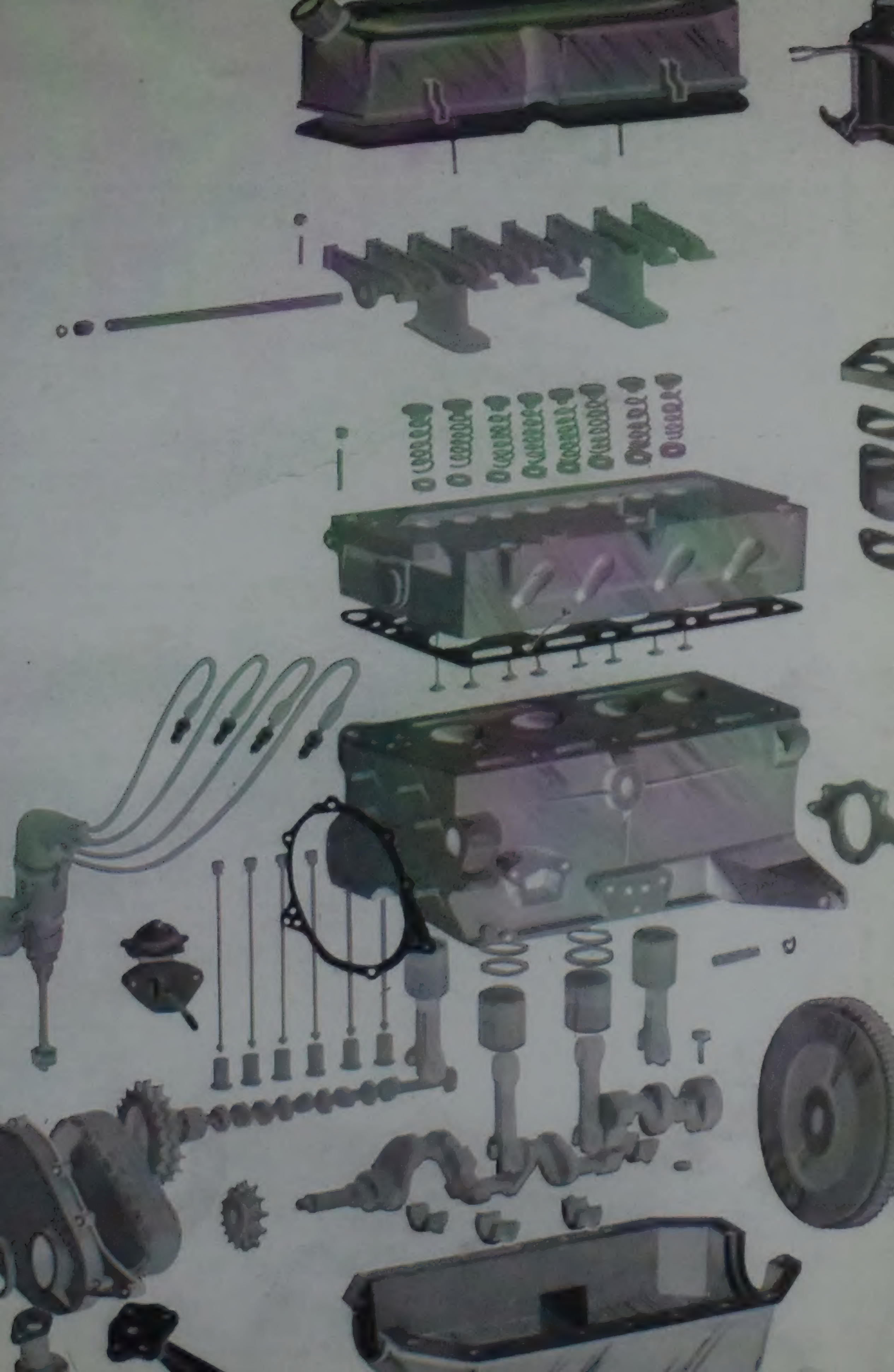
SIRO La leggenda continua Un CD con i cori inglesi e della curva sud del Milan, e i pezzi musicali originali di Sergio Messina, Royalize, Sigma Tibet e Adrian Sheerwood, CD audio di 47 minuti con opuscolo 16 pp. a colori, E. 10,70 • **LE PAROLE E LA LOTTA ARMATA - Storia vissuta e sinistra militante in Italia, Germania e Svizzera,** a cura di Primo Moroni, IG Rote Fabrik, Konzeptbüro. Nel 1997 a Zurigo si radunarono per alcuni giorni gli esponenti di alcune delle più importanti organizzazioni "armate" degli anni settanta-ottanta per discutere senza veli con altri militanti e storici della loro esperienza, pp. 224, E. 14,46.

VIDEO

VIDEOZINE CYBERPUNK vol. I. Videocassetta con libretto, antologia di futuri contemporanei, 80 minuti, E. 12,91 • **VIDEOZINE CYBERPUNK vol. II**. Vhs con libretto, 90 minuti, E. 12,91 • **COMMISSIONER OF SEWERS** William Burroughs. Un film di Klaus Maeck VHS 60 minuti di documenti su zio Bill, altamente professionali, E. 18,08 • **DECODER IL FILM** Klaus Maeck. Un film cult e profetico. Con F.M. Einheit, G.P-Orridge, Christiana F., W.S. Burroughs. Musica: Einstürzende Neubauten, Soft Cell, The The, 58 minuti, E. 18,08 • **VHS VIRUS** Documento storico, girato in 16 mm. nel 1982 all'interno della casa occupata dai punk di via Correggio 18 a Milano. Strettamente collegato al romanzo *Costretti a Sanguinare*, durata 30 minuti, E. 12,91 • **Richard Stallman FILOSOFIA E PRATICA DEL PROGETTO GNU**. Il programmatore Richard Stallman, in questa conferenza tenuta in Cox 18, ci spiega del Progetto Gnu, uno dei più interessanti esperimenti per la produzione di una vera, concreta alternativa al monopolio delle grandi aziende del software, durata 90 min., E. 12,91.• **SOLO LIMONI**. Giacomo Verde e Lello Voce. Videodocumentazione poetica dei fatti di Genova nel luglio 2001. VHS accompagnata da un libro che offre i contributi originali di molti tra gli autori della nuova letteratura italiana, durata 46 minuti, E. 14,46.

RIVISTE

DECODER 12 Speciale privacy, Primo Moroni, Controcultura informatica, Richard Stallman, e-money, Push technology, Pat Cadigan, Traveller, Letteratura Avant Pop, Mike Davis, Reclaim the Street, Ecstasy, abbiamo raggiunto p. 1000 di controculture radicali e libertarie, pp. 72, E. 4,65 • **DECODER 11** Speciale J.G. Ballard, Hakim Bey, Pierre Lévy, Neal Stephenson, Cyberfemminismo, Jungle, Helena Velena, pp. 96, E. 4,65 • **DECODER RACCOLTA N. 5/6** Ristampa dei numeri della rivista internazionale underground usciti prima della fine degli anni Ottanta, pp. 192, E. 12,91, sono ancora disponibili le ultimissime copie dei nn. 8/9/10 • **FIKA FUTURA Rivista di secrezioni acide cyberfemministe & queer** n. 2 Helena Velena, Surgical Beauty, Speed Demon, Sbarbies Adventures, pp. 36, E. 3,10 • **FIKA FUTURA Rivista di secrezioni acide cyberfemministe & queer** n. 1 VNS Matrix, Rete Femminile Singolare, Katy Acker, Feti in faccia (Fumetto in 3d), Autodifesa femminile, Frutta interattiva, Linda Dement, pp. 32, E. 3,10 • **PSYCHOattiva Rivista di vita e cultura psichedelica**. n. 1, pp. 112, E. 6,20 • **PSYCHOattiva Rivista di vita e cultura psichedelica**. n. 2, pp. 96, E. 6,20. • **ACOMA Rivista Internazionale di Studi Nordamericani** n. 21, pp.160, E. 9,30. • **ACOMA Rivista Internazionale di Studi Nordamericani** n. 22, pp.128, E. 9,30.



Questo libro è un'indagine sui motori di ricerca, sulle loro caratteristiche, limiti ed evoluzioni, proprio perché la ricerca dell'informazione in Rete e le tecniche per rendere l'informazione più visibile nel Web hanno assunto negli ultimi tempi un'importanza sempre più grande.
I motori di ricerca raramente però sono messi in discussione o guardati al microscopio con occhi disincantati. Cosa sono gli spider? Cosa realmente vedono? Cosa cercano i motori di ricerca? Esiste un kit di sopravvivenza per i profani e gli esperti che si affacciano su questo magma fluttuante di informazioni? Esiste un reale pericolo di "censura" dell'informazione, o c'è pari visibilità per tutti su Internet? L'informazione è in effetti "universalmente accessibile"?
Nello spiegare in modo dettagliato il funzionamento dei motori di ricerca e delle web directory, le modalità di ricerca e tecniche quali il "top ranking", per apparire fra le prime posizioni nei risultati di ricerca, gli autori concretamente ci parlano di diritto al sapere, puntando direttamente alla nuova "frontiera dell'accessibilità", nodo cruciale nel futuro del World Wide Web.

CLAUDIO PARRINI (1963), networker e pittore, è membro del gruppo Stranonetwork e collabora con UnDo.Net. Il suo lavoro si sviluppa attraverso la realizzazione di opere sia di natura comunitaria sia individuale: progetti in rete, laboratori, quadri e scritti.
FERRY BYTE (1967), hacktivist, è impegnato da tempo sul versante delle mobilitazioni elettroniche dell'autorganizzazione digitale (i netstrike) e nella promozione di forme accessibili dell'informazione in Internet. È fondatore di Stranonetwork e socio di Isole nella Rete.

CYBERPUNKLINE

Euro 9,30

ISBN 88-86926-75-8
9 788886 926751